

# HARDY SPACE THEORY ON SPACES OF HOMOGENEOUS TYPE VIA ORTHONORMAL WAVELET BASES

YONGSHENG HAN, JI LI AND LESLEY A. WARD

**ABSTRACT.** In this paper, using the remarkable orthonormal wavelet basis constructed recently by Auscher and Hytönen, we establish the theory of product Hardy spaces on spaces  $\tilde{X} = X_1 \times X_2 \times \cdots \times X_n$ , where each factor  $X_i$  is a space of homogeneous type in the sense of Coifman and Weiss. The main tool we develop is the Littlewood–Paley theory on  $\tilde{X}$ , which in turn is a consequence of a corresponding theory on each factor space. We define the square function for this theory in terms of the wavelet coefficients. The Hardy space theory developed in this paper includes product  $H^p$ , the dual  $\text{CMO}^p$  of  $H^p$  with the special case  $\text{BMO} = \text{CMO}^1$ , and the predual  $\text{VMO}$  of  $H^1$ . We also use the wavelet expansion to establish the Calderón–Zygmund decomposition for product  $H^p$ , and deduce an interpolation theorem. We make no additional assumptions on the quasi-metric or the doubling measure for each factor space, and thus we extend to the full generality of product spaces of homogeneous type the aspects of both one-parameter and multiparameter theory involving the Littlewood–Paley theory and function spaces. Moreover, our methods would be expected to be a powerful tool for developing wavelet analysis on spaces of homogeneous type.

## CONTENTS

1. Introduction	2
2. Preliminaries	7
2.1. Systems of dyadic cubes in a geometrically doubling metric space	8
2.2. Orthonormal wavelet basis and wavelet expansion	8
3. Test functions, distributions, and wavelet reproducing formula	10
3.1. One-parameter test functions, distributions, and wavelet reproducing formula	10
3.2. Product test functions, distributions, and wavelet reproducing formula	25
4. Littlewood–Paley square functions and Plancherel–Pólya inequalities	28
4.1. One-parameter square functions via wavelets, and Plancherel–Pólya inequalities	28
4.2. Product square functions via wavelets, and Plancherel–Pólya inequalities	39
5. Product $H^p$ , $\text{CMO}^p$ , $\text{BMO}$ and $\text{VMO}$ , and duality	40
6. Calderón–Zygmund decomposition and interpolation on Hardy spaces	45
References	49

2010 *Mathematics Subject Classification.* Primary 42B35; Secondary 43A85, 42B25, 42B30.

*Key words and phrases.* Spaces of homogeneous type, orthonormal basis, test function space, distributions, Calderón reproducing formula, wavelet expansion, product Hardy space, Carleson measure space,  $\text{BMO}$ ,  $\text{VMO}$ , duality.

The second and third authors are supported by the Australian Research Council under Grant No. arc-dp120100399. The second author was also supported by the NNSF of China Grant No. 11001275, by China Postdoctoral Science Foundation funded project Grant No. 201104383, and by the Fundamental Research Funds for the Central Universities, Grant No. 11lgpy56. Parts of this paper were written while the second author was a member of the Department of Mathematics, Sun Yat-sen University, Guangzhou, 510275, P.R. China.

## 1. INTRODUCTION

We work on wavelet analysis in the setting of product spaces of homogeneous type in the sense of Coifman and Weiss [CW1], where each factor is of the form  $(X, d, \mu)$  with  $d$  a quasi-metric and  $\mu$  a doubling measure. We make no additional assumptions on  $d$  or  $\mu$ . After recalling the systems of dyadic cubes of Hytönen and Kairema [HK] and the orthonormal wavelet basis of Auscher and Hytönen [AH], we define an appropriate class of test functions and the induced class of distributions on spaces of homogeneous type. We prove that the Auscher–Hytönen wavelets are test functions, and that the Auscher–Hytönen reproducing formula for  $L^p$  also holds for our test functions and distributions. We show that the kernels of certain wavelet operators  $D_k$  defined in terms of these wavelets satisfy decay and smoothness conditions similar to those of our test functions. These facts play a crucial role in our development of the Littlewood–Paley theory and function spaces, later in our paper.

We define the discrete Littlewood–Paley square function via the Auscher–Hytönen wavelet coefficients. In order to establish its  $L^p$ -boundedness, we also introduce a different, *continuous* Littlewood–Paley square function defined in terms of the wavelet operators  $D_k$ . We prove that the discrete and continuous square functions have equivalent norms, by first establishing some inequalities of Plancherel–Pólya type. We develop this Littlewood–Paley theory first in the one-parameter setting, and then for product spaces.

For  $p$  in a range that depends on the upper dimensions of the spaces  $X_1$  and  $X_2$  and strictly includes the range  $1 \leq p < \infty$ , we define the product Hardy space  $H^p(X_1 \times X_2)$  as the class of distributions whose discrete Littlewood–Paley square functions are in  $L^p(X_1 \times X_2)$ . (Here we write only two factors, for simplicity, but our results extend to  $n$  factors.) For  $p$  in this range with  $p \leq 1$ , we define the Carleson measure space  $\text{CMO}^p(X_1 \times X_2)$  via the Auscher–Hytönen wavelet coefficients, as a subset of our space of distributions, and prove the duality  $(H^p(X_1 \times X_2))' = \text{CMO}^p(X_1 \times X_2)$  by means of sequence spaces that form discrete analogues of these spaces. This duality result includes the special case  $(H^1(X_1 \times X_2))' = \text{BMO}(X_1 \times X_2)$ . We define the space  $\text{VMO}(X_1 \times X_2)$  of functions of vanishing mean oscillation, also in terms of the Auscher–Hytönen wavelet coefficients, and prove the duality  $(\text{VMO}(X_1 \times X_2))' = H^1(X_1 \times X_2)$  by adapting an argument of Lacey–Terwilleger–Wick [LTW]. Using the wavelet expansion, we establish the Calderón–Zygmund decomposition for functions in our Hardy spaces  $H^p(X_1 \times X_2)$ , again for a suitable range of  $p$  that strictly includes  $1 \leq p < \infty$ . As a consequence, we deduce an interpolation theorem for linear operators from these product Hardy spaces to Lebesgue spaces on  $X_1 \times X_2$ .

We now set our work in context. As Meyer remarked in his preface to [DH], “*One is amazed by the dramatic changes that occurred in analysis during the twentieth century. In the 1930s complex methods and Fourier series played a seminal role. After many improvements, mostly achieved by the Calderón–Zygmund school, the action takes place today on spaces of homogeneous type. No group structure is available, the Fourier transform is missing, but a version of harmonic analysis is still present. Indeed the geometry is conducting the analysis.*” Spaces of homogeneous type were introduced by Coifman and Weiss in the early 1970s, in [CW1]. We say that  $(X, d, \mu)$  is a *space of homogeneous type* in the sense of Coifman and Weiss if  $d$  is a quasi-metric on  $X$  and  $\mu$  is a nonzero measure satisfying the doubling condition. A *quasi-metric*  $d$  on a set  $X$  is a function  $d : X \times X \rightarrow [0, \infty)$  satisfying (i)

$d(x, y) = d(y, x) \geq 0$  for all  $x, y \in X$ ; (ii)  $d(x, y) = 0$  if and only if  $x = y$ ; and (iii) the *quasi-triangle inequality*: there is a constant  $A_0 \in [1, \infty)$  such that for all  $x, y, z \in X$ ,

$$(1.1) \quad d(x, y) \leq A_0[d(x, z) + d(z, y)].$$

We define the quasi-metric ball by  $B(x, r) := \{y \in X : d(x, y) < r\}$  for  $x \in X$  and  $r > 0$ . Note that the quasi-metric, in contrast to a metric, may not be Hölder regular and quasi-metric balls may not be open. We say that a nonzero measure  $\mu$  satisfies the *doubling condition* if there is a constant  $C_\mu$  such that for all  $x \in X$  and  $r > 0$ ,

$$(1.2) \quad \mu(B(x, 2r)) \leq C_\mu \mu(B(x, r)) < \infty.$$

We point out that the doubling condition (1.2) implies that there exist positive constants  $\omega$  (the *upper dimension* of  $\mu$ ) and  $C$  such that for all  $x \in X$ ,  $\lambda \geq 1$  and  $r > 0$ ,

$$(1.3) \quad \mu(B(x, \lambda r)) \leq C \lambda^\omega \mu(B(x, r)).$$

Spaces of homogeneous type include many special spaces in analysis and have many applications in the theory of singular integrals and function spaces; see [CW2, NS1, NS2] for more detail. For instance, Coifman and Weiss introduced the atomic Hardy space on  $(X, d, \mu)$  and proved that if  $T$  is a Calderón–Zygmund singular integral operator and is bounded on  $L^2(X)$ , then  $T$  is bounded from  $H^p(X)$  to  $L^p(X)$  for some  $p \leq 1$ .

However, for some applications, additional assumptions were imposed on these general spaces of homogeneous type, because as noted above the original quasi-metric  $d$  may have no regularity and quasi-metric balls, even Borel sets, may not be open. For example, to establish the maximal function characterization of the Hardy space introduced by Coifman and Weiss, Macías and Segovia proved in [MS] that one can replace the quasi-metric  $d$  by another quasi-metric  $d'$  on  $X$  such that the topologies induced on  $X$  by  $d$  and  $d'$  coincide, and  $d'$  has the following regularity property:

$$(1.4) \quad |d'(x, y) - d'(x', y)| \leq C_0 d'(x, x')^\theta [d'(x, y) + d'(x', y)]^{1-\theta}$$

for some constant  $C_0$ , some regularity exponent  $\theta \in (0, 1)$ , and for all  $x, x', y \in X$ . Moreover, if quasi-metric balls are defined by this new quasi-metric  $d'$ , that is,  $B'(x, r) := \{y \in X : d'(x, y) < r\}$  for  $r > 0$ , then the measure  $\mu$  satisfies the following property:

$$(1.5) \quad \mu(B'(x, r)) \sim r.$$

Note that property (1.5) is much stronger than the doubling condition. Macías and Segovia established the maximal function characterization for Hardy spaces  $H^p(X)$  with  $(1 + \theta)^{-1} < p \leq 1$ , on spaces of homogeneous type  $(X, d', \mu)$  that satisfy the regularity condition (1.4) on the metric  $d'$  and property (1.5) on the measure  $\mu$ ; see [MS].

A fundamental result for these spaces  $(X, d', \mu)$  is the  $T(b)$  theorem of David–Journé–Semmes [DJS]. The crucial tool in the proof of the  $T(b)$  theorem is the existence of a suitable approximation to the identity. The construction of such an approximation to the identity is due to Coifman. More precisely, take a smooth function  $h$  defined on  $[0, \infty)$ , equal to 1 on  $[1, 2]$ , and equal to 0 on  $[0, 1/2]$  and on  $[4, \infty)$ . Let  $T_k$  be the operator with kernel  $2^k h(2^k d'(x, y))$ . The property (1.5) of the measure  $\mu$  implies that  $C^{-1} \leq T_k(1) \leq C$  for some  $C$  with  $0 < C < \infty$ . Let  $M_k$  and  $W_k$  be the operators of multiplication by  $1/T_k(1)$  and  $\{T_k[1/T_k(1)]\}^{-1}$ , respectively, and let  $S_k := M_k T_k W_k T_k M_k$ . It is clear that the regularity

property (1.4) on the metric  $d'$  and property (1.5) on the measure  $\mu$  imply that the kernel  $S_k(x, y)$  of  $S_k$  satisfies the following conditions: for some constants  $C > 0$  and  $\varepsilon > 0$ ,

- (i)  $S_k(x, y) = 0$  for  $d'(x, y) \geq C2^{-k}$ , and  $\|S_k\|_\infty \leq C2^k$ ,
- (ii)  $|S_k(x, y) - S_k(x', y)| \leq C2^{k(1+\varepsilon)}d'(x, x')^\varepsilon$ ,
- (iii)  $|S_k(x, y) - S_k(x, y')| \leq C2^{k(1+\varepsilon)}d'(y, y')^\varepsilon$ , and
- (iv)  $\int_X S_k(x, y) d\mu(y) = 1 = \int_X S_k(x, y) d\mu(x)$ .

Let  $D_k := S_{k+1} - S_k$ . In [DJS], the Littlewood–Paley theory for  $L^p(X)$ ,  $1 < p < \infty$ , was established; namely, if  $\mu(X) = \infty$  and  $\mu(B(x, r)) > 0$  for all  $x \in X$  and  $r > 0$ , then for each  $p$  with  $1 < p < \infty$  there exists a positive constant  $C_p$  such that

$$C_p^{-1}\|f\|_p \leq \left\| \left\{ \sum_k |D_k(f)|^2 \right\}^{\frac{1}{2}} \right\|_p \leq C_p\|f\|_p.$$

The above estimates were the key tool for proving the  $T(1)$  theorem on  $(X, d', \mu)$ ; see [DJS] for more detail. Later, the Calderón reproducing formula, test function spaces and distributions, the Littlewood–Paley theory, and function spaces on  $(X, d', \mu)$  were developed in [H1], [HS] and [H2]. However, in those works wavelet bases were replaced by frames, which in many applications offer the same service; see [DH] for more details.

In [NS1], Nagel and Stein developed the product  $L^p$  ( $1 < p < \infty$ ) theory in the setting of the Carnot–Carathéodory spaces formed by vector fields satisfying Hörmander’s finite rank condition. The Carnot–Carathéodory spaces studied in [NS1] are spaces of homogeneous type with a smooth quasi-metric  $d$  and a measure  $\mu$  satisfying the conditions  $\mu(B(x, sr)) \sim s^{m+2}\mu(B(x, r))$  for  $s \geq 1$  and  $\mu(B(x, sr)) \sim s^4\mu(B(x, r))$  for  $s \leq 1$ . These conditions on the measure are weaker than property (1.5) but are still stronger than the original doubling condition (1.2). In [HMY], motivated by the work of Nagel and Stein, Hardy spaces were developed on spaces of homogeneous type with a regular quasi-metric and a measure satisfying the above conditions. Recently, in [HLL2], it was observed that Coifman’s construction of an approximation to the identity still works on spaces of homogeneous type  $(X, d, \mu)$  where the quasi-metric  $d$  satisfies the Hölder regularity property (1.4) but the measure  $\mu$  only needs to be doubling. Specifically, the kernel  $S_k(x, y)$  of the approximation to the identity  $S_k$  satisfies the following conditions: there exist constants  $C > 0$  and  $\theta > 0$  such that for all  $k \in \mathbb{Z}$  and all  $x, x', y, y' \in X$ ,

- (i)  $S_k(x, y) = 0$  for  $d(x, y) \geq C2^{-k}$ , and  $|S_k(x, y)| \leq C \frac{1}{V_{2^{-k}}(x) + V_{2^{-k}}(y)}$ ,
- (ii)  $|S_k(x, y) - S_k(x', y)| \leq C2^{k\theta}d(x, x')^\theta \frac{1}{V_{2^{-k}}(x) + V_{2^{-k}}(y)}$ ,
- (iii) property (ii) also holds with  $x$  and  $y$  interchanged,
- (iv)  $|[S_k(x, y) - S_k(x, y')] - [S_k(x', y) - S_k(x', y')]|$   
 $\leq C2^{2k\theta}d(x, x')^\theta d(y, y')^\theta \frac{1}{V_{2^{-k}}(x) + V_{2^{-k}}(y)},$  and
- (v)  $\int_X S_k(x, y) d\mu(y) = 1 = \int_X S_k(x, y) d\mu(x)$ ,

where  $V_r(x) := \mu(B(x, r))$ .

For spaces of homogeneous type  $(X, d, \mu)$  with some additional assumptions, the one-parameter and product Hardy spaces were developed in [HMY] and [HLL2], respectively.

A natural question arises: can one develop the theory of the spaces  $H^p$  and BMO on spaces of homogeneous type in the sense of Coifman and Weiss, with only the original quasi-metric  $d$  and a doubling measure  $\mu$ ?

Recently, Auscher and Hytönen constructed an orthonormal wavelet basis with Hölder regularity and exponential decay for spaces of homogeneous type [AH]. This result is remarkable since there are no additional assumptions other than those defining spaces of homogeneous type in the sense of Coifman and Weiss.

Auscher and Hytönen's orthonormal wavelet bases open the door for developing wavelet analysis on spaces of homogeneous type in the sense of Coifman and Weiss. Motivated by Auscher and Hytönen's work, the purpose of the current paper is to answer the above question. Specifically, we will employ a unified approach to establish a product Hardy space theory on  $\tilde{X} = X_1 \times X_2 \times \cdots \times X_n$ , where each factor is a space of homogeneous type in the sense of Coifman and Weiss. It was well known that any analysis of the product Hardy space on  $\tilde{X} = X_1 \times X_2 \times \cdots \times X_n$  must be based, to start with, on a formulation on each factor  $X_j$ . The Hardy space on  $X_j$  is then defined by developing the Littlewood–Paley theory on  $X_j$ . Our approach includes the following five steps.

**1. Introduce the spaces of test functions and distributions.** In the classical case, the relevant spaces of test functions and distributions are just Schwartz test functions and the class of tempered distributions. In order to study the Calderón reproducing formula associated with the  $T(b)$  theorem, the new test function and distribution spaces were first introduced on Euclidean spaces in [H1], and on spaces of homogeneous type, where the quasi-metric  $d$  satisfies the Hölder regularity condition (1.4) and the measure  $\mu$  satisfies condition (1.5), in [HS]. See [HMY] and [HLL2], respectively, for spaces of test functions and distributions on spaces of homogeneous type with additional assumptions. In this paper, we will introduce test functions and distributions on spaces of homogeneous type in the sense of Coifman and Weiss. These spaces include all those considered previously.

**2. Establish the wavelet reproducing formula on test functions and on distributions.** The classical Calderón reproducing formula was first used by Calderón in [C]. Such a reproducing formula is a powerful tool, particularly in the theory of wavelet analysis. See [M1]. Using Coifman's decomposition of the identity operator, as mentioned above, David, Journé and Semmes [DJS] gave a Calderón-type reproducing formula which was a key tool in proving the  $T(b)$  theorem on  $\mathbb{R}^n$  and the  $T(1)$  theorem on spaces of homogeneous type with the conditions (1.4) and (1.5). See [HMY] and [HLL2] for the continuous and discrete Calderón reproducing formulas on spaces of homogeneous type with additional assumptions. As mentioned above, Auscher and Hytönen established a wavelet expansion on  $L^2(X)$  (and on  $L^p(X)$ ,  $1 < p < \infty$ ). For our purposes, we will show that the wavelet expansion constructed in [AH] also converges in both the test function and distribution spaces.

As Meyer pointed out in [M1], *“The wavelet bases are universally applicable: ‘everything that comes to hand’, whether function or distribution, is the sum of a wavelet series and, contrary to what happens with Fourier series, the coefficients of the wavelet series translate the properties of the function or distribution simply, precisely and faithfully.”* In particular,

our results provide such wavelet expansions for test functions and for distributions, and are used below to introduce square functions and develop the Littlewood–Paley theory.

**3. Develop the Littlewood–Paley theory.** Based on the wavelet expansion provided in [AH], one can formally introduce two kinds of square functions, namely, the discrete version defined in terms of wavelet coefficients and the continuous version defined via wavelet operators  $D_k$  (different from the operators  $D_k = S_{k+1} - S_k$  mentioned above). To show that the  $L^p$  norms of these square functions are equivalent, for a suitable range of  $p$ , we need a Plancherel–Pólya inequality. The classical Plancherel–Pólya inequality says that the  $L^p$  norm of a function  $f$  whose Fourier transform has compact support is equivalent to the  $\ell^p$  norm of the restriction of  $f$  to an appropriate lattice. This kind of inequality was first proved in [H2] on spaces of homogeneous type with the conditions (1.4) and (1.5), and in [HMY] and [HLL2], respectively, for the one-parameter and multiparameter cases with some additional assumptions. As a consequence of our Plancherel–Pólya type inequalities, the Hardy space on spaces of homogeneous type in the sense of Coifman and Weiss is well defined. In particular, as in the classical case,  $H^p = L^p$  for  $1 < p < \infty$ .

**4. Introduce the generalized Carleson measure space.** It is well known that in the classical one-parameter case, the space BMO, as the dual of  $H^1$ , can be characterized by Carleson measures. Moreover, in [CF1] Chang and Fefferman proved that the dual of product  $H^1$  is characterized by product Carleson measures. The generalized Carleson measure space  $\text{CMO}^p$ , as the dual of the product  $H^p$ , was introduced in [HLL1] and [HLL2] on spaces of homogeneous type with some additional assumptions. In the current paper, working in the setting of spaces of homogeneous type in the sense of Coifman and Weiss with no additional assumptions, we introduce  $\text{CMO}^p$  in terms of wavelet coefficients, and prove that  $\text{CMO}^p$  is the dual of  $H^p$ . In particular,  $\text{CMO}^1 = \text{BMO}$  is the dual of  $H^1$ . Moreover, we also introduce the space VMO and show that VMO is the predual of  $H^1$ .

**5. Establish the Calderón–Zygmund decomposition.** The classical Calderón–Zygmund decomposition played a crucial role in developing Calderón–Zygmund operator theory. This decomposition has many applications in harmonic analysis and partial differential equations. Such a decomposition for product Euclidean spaces was first provided by Chang and Fefferman in [CF2]. The main tool used in [CF2] is the atomic decomposition. In the current paper, applying the wavelet expansion constructed in [AH], we establish the Calderón–Zygmund decomposition on product  $H^p$  on spaces of homogeneous type with no additional assumptions. As a consequence, we obtain the interpolation of operators that are bounded from Hardy spaces to Lebesgue spaces, and of operators that are bounded on Hardy spaces.

We note that in the original work on extending the Calderón–Zygmund theory to spaces of homogeneous type  $(X, d', \mu)$ , the philosophy was as follows: Coifman constructed the approximations to the identity  $S_k$ , which were used in [DJS] to define the continuous square function and to establish the Littlewood–Paley theory. Later the discrete Calderón reproducing formula was introduced and the Littlewood–Paley theory for the classical function spaces were established in [H1] and [HS], respectively. By contrast, in our setting of  $(X, d, \mu)$  with the original quasi-metric  $d$ , we begin with the discrete wavelet reproducing formula (Theorem 3.4) and define the discrete square function  $S(f)$  in terms of wavelet coefficients (Definition 4.1).

However, there does not seem to be a direct proof of the Littlewood–Paley theory for  $S(f)$ . The question then is: how to find a continuous version of the square function? We introduce a new continuous square function  $S_c(f)$  (Definition 4.2), via certain wavelet operators  $D_k$  that are expressed in terms of the Auscher–Hytönen wavelets (Lemma 3.6). We prove that  $\|S_c(f)\|_p \sim \|f\|_p$  for  $1 < p < \infty$  (Theorem 4.4), and that  $\|S(f)\|_p \sim \|S_c(f)\|_p$  both for  $1 < p < \infty$  and moreover for an additional range of  $p \leq 1$  depending on the upper dimensions  $\omega_i$  of the factor spaces  $X_i$  and on the Hölder regularity exponents  $\eta_i$  of the wavelets (Theorem 4.3).

We remark that in this paper we concentrate on the product case. As Nagel and Stein observed in [NS1], “*Any product theory tends to be burdened with notational complexities.*” For notational simplicity, we have written our results and proofs for the case of two parameters. However, our methods also establish the corresponding results for the product case with  $k$  factors, for  $k \in \mathbb{N}$ . We also point out that these results extend related previous results from [DJS, H1, H2, HLL1, HLL2, HMY, HS] and the references therein. In those papers either extra assumptions are made on the quasi-metric and the measure, or the product case is not covered, or both.

The paper is organized as follows. In Section 2 we briefly recall the systems of dyadic cubes from [HK] and the orthonormal bases from [AH] on spaces of homogeneous type in the sense of Coifman and Weiss. In Section 3 we introduce the one-parameter and product test functions in Definitions 3.1 and 3.9, respectively, together with the induced classes of distributions. The main result in this section is Theorem 3.4, which gives the wavelet reproducing formula for test functions. In Section 4, the Littlewood–Paley square functions in terms of the wavelet coefficients and of the wavelet operators are given in Definitions 4.1 and 4.2, respectively. The two main results here are Theorems 4.3 and 4.4. Theorem 4.3 gives the Littlewood–Paley theory, including the norm equivalence of the discrete and continuous Littlewood–Paley square functions. Theorem 4.4 gives the Plancherel–Pólya inequalities, which are the main tool in proving Theorem 4.3. The product  $H^p$ ,  $\text{CMO}^p$ ,  $\text{BMO}$  and  $\text{VMO}$  spaces are defined in Section 5 via the orthonormal wavelet basis. We use the Plancherel–Pólya inequalities again to show that these function spaces are well defined. The duality results are given in Theorem 5.3 for  $(H^p)' = \text{CMO}^p$  and in Theorem 5.10 for  $(\text{VMO})' = H^1$ . Finally, in Section 6 we prove the Calderón–Zygmund decomposition and the interpolation theorem for Hardy spaces in Theorems 6.1 and 6.2, respectively.

## 2. PRELIMINARIES

We are interested in establishing the Hardy space theory on spaces  $\tilde{X} = X_1 \times X_2 \times \cdots \times X_n$ . Each factor is a space of homogeneous type in the sense of Coifman and Weiss. We will first need to develop a Littlewood–Paley theory for each factor  $X_i$ ,  $1 \leq i \leq n$ , and then pass to the corresponding product theory. In this paper, we always assume that  $\mu(X_i) = \infty$  and that  $\mu(B(x, r)) > 0$  for all  $r > 0$  and  $x \in X_i$ , for  $1 \leq i \leq n$ . As usual,  $C$  denotes a constant that is independent of the essential variables, and that may differ from line to line.

In this section we recall the systems of dyadic cubes, in a geometrically doubling metric space, constructed by Hytönen and Kairema [HK]; and the orthonormal wavelet basis, on spaces of homogeneous type, constructed by Auscher and Hytönen [AH, AH2]. See also [HK],

[AH] and the references therein for the history and applications of various versions of dyadic cubes.

**2.1. Systems of dyadic cubes in a geometrically doubling metric space.** Let  $d$  be a quasi-metric (defined in the Introduction) on a set  $X$ . The quasi-metric space  $(X, d)$  is assumed to have the following *geometric doubling property*: there exists a positive integer  $A_1 \in \mathbb{N}$  such that for each  $x \in X$  and each  $r > 0$ , the ball  $B(x, r) := \{y \in X : d(y, x) < r\}$  can be covered by at most  $A_1$  balls  $B(x_i, r/2)$ . It is shown in [CW1] that spaces of homogeneous type have the geometric doubling property.

As usual, a set  $\Omega \subset X$  is *open* if for every  $x \in \Omega$  there exists  $\varepsilon > 0$  such that  $B(x, \varepsilon) \subset \Omega$ , and a set is *closed* if its complement is open.

**Theorem 2.1** ([HK] Theorem 2.2). *Suppose that constants  $0 < c_0 \leq C_0 < \infty$  and  $\delta \in (0, 1)$  satisfy*

$$(2.1) \quad 12A_0^3 C_0 \delta \leq c_0.$$

*Given a set of points  $\{z_\alpha^k\}_\alpha$ ,  $\alpha \in \mathcal{A}_k$ , for every  $k \in \mathbb{Z}$ , with the properties that*

$$(2.2) \quad d(z_\alpha^k, z_\beta^k) \geq c_0 \delta^k \ (\alpha \neq \beta), \quad \min_\alpha d(x, z_\alpha^k) < C_0 \delta^k, \quad \text{for all } x \in X,$$

*we can construct families of sets  $\tilde{Q}_\alpha^k \subseteq Q_\alpha^k \subseteq \overline{Q}_\alpha^k$  (called open, half-open and closed dyadic cubes), such that:*

$$(2.3) \quad \tilde{Q}_\alpha^k \text{ and } \overline{Q}_\alpha^k \text{ are the interior and closure of } Q_\alpha^k, \text{ respectively;}$$

$$(2.4) \quad \text{if } \ell \geq k, \text{ then either } Q_\beta^\ell \subseteq Q_\alpha^k \text{ or } Q_\alpha^k \cap Q_\beta^\ell = \emptyset;$$

$$(2.5) \quad X = \bigcup_{\alpha} Q_\alpha^k \text{ (disjoint union) for all } k \in \mathbb{Z};$$

$$(2.6) \quad B(z_\alpha^k, c_1 \delta^k) \subseteq Q_\alpha^k \subseteq B(z_\alpha^k, C_1 \delta^k), \text{ where } c_1 := (3A_0^2)^{-1} c_0 \text{ and } C_1 := 2A_0 C_0;$$

$$(2.7) \quad \text{if } \ell \geq k \text{ and } Q_\beta^\ell \subseteq Q_\alpha^k, \text{ then } B(z_\beta^\ell, C_1 \delta^\ell) \subseteq B(z_\alpha^k, C_1 \delta^k).$$

*The open and closed cubes  $\tilde{Q}_\alpha^k$  and  $\overline{Q}_\alpha^k$  depend only on the points  $z_\beta^\ell$  for  $\ell \geq k$ . The half-open cubes  $Q_\alpha^k$  depend on  $z_\beta^\ell$  for  $\ell \geq \min(k, k_0)$ , where  $k_0 \in \mathbb{Z}$  is a preassigned number entering the construction.*

**2.2. Orthonormal wavelet basis and wavelet expansion.** In this subsection, we recall the orthonormal basis and wavelet expansion in  $L^2(X)$  which were recently constructed by Auscher and Hytönen [AH]. To state their result, we must first recall the set  $\{x_\alpha^k\}$  of *reference dyadic points* as follows. Let  $\delta$  be a fixed small positive parameter (for example, as noted in Section 2.2 of [AH], it suffices to take  $\delta \leq 10^{-3} A_0^{-10}$ ). For  $k = 0$ , let  $\mathcal{X}^0 := \{x_\alpha^0\}_\alpha$  be a maximal collection of 1-separated points in  $X$ . Inductively, for  $k \in \mathbb{Z}_+$ , let  $\mathcal{X}^k := \{x_\alpha^k\} \supseteq \mathcal{X}^{k-1}$  and  $\mathcal{X}^{-k} := \{x_\alpha^{-k}\} \subseteq \mathcal{X}^{-(k-1)}$  be  $\delta^k$ - and  $\delta^{-k}$ -separated collections in  $\mathcal{X}^{k-1}$  and  $\mathcal{X}^{-(k-1)}$ , respectively.

Lemma 2.1 in [AH] shows that, for all  $k \in \mathbb{Z}$  and  $x \in X$ , the reference dyadic points satisfy

$$(2.8) \quad d(x_\alpha^k, x_\beta^k) \geq \delta^k \ (\alpha \neq \beta), \quad d(x, \mathcal{X}^k) = \min_\alpha d(x, x_\alpha^k) < 2A_0 \delta^k.$$



Also, taking  $c_0 := 1$ ,  $C_0 := 2A_0$  and  $\delta \leq 10^{-3}A_0^{-10}$ , we see that  $c_0$ ,  $C_0$  and  $\delta$  satisfy (2.1) in Theorem 2.1. Therefore we may apply Hytönen and Kairema's construction (Theorem 2.1), with the reference dyadic points  $\{x_\alpha^k\}_{k \in \mathbb{Z}, \alpha \in \mathcal{X}^k}$  playing the role of the points  $\{z_\alpha^k\}_{k \in \mathbb{Z}, \alpha \in \mathcal{A}_k}$ , to conclude that there exists a set of half-open dyadic cubes

$$\{Q_\alpha^k\}_{k \in \mathbb{Z}, \alpha \in \mathcal{X}^k}$$

associated with the reference dyadic points  $\{x_\alpha^k\}_{k \in \mathbb{Z}, \alpha \in \mathcal{X}^k}$ . We call the reference dyadic point  $x_\alpha^k$  the *center* of the dyadic cube  $Q_\alpha^k$ . We also identify with  $\mathcal{X}^k$  the set of indices  $\alpha$  corresponding to  $x_\alpha^k \in \mathcal{X}^k$ .

Note that  $\mathcal{X}^k \subseteq \mathcal{X}^{k+1}$  for  $k \in \mathbb{Z}$ , so that every  $x_\alpha^k$  is also a point of the form  $x_\beta^{k+1}$ . We denote  $\mathcal{Y}^k := \mathcal{X}^{k+1} \setminus \mathcal{X}^k$ , and relabel the points  $\{x_\alpha^k\}_\alpha$  that belong to  $\mathcal{Y}^k$  as  $\{y_\alpha^k\}_\alpha$ .

We now recall the orthonormal wavelet basis of  $L^2(X)$  constructed by Auscher and Hytönen.

**Theorem 2.2** ([AH] Theorem 7.1). *Let  $(X, d, \mu)$  be a space of homogeneous type with quasi-triangle constant  $A_0$ , and let*

$$(2.9) \quad a := (1 + 2 \log_2 A_0)^{-1}.$$

*There exists an orthonormal wavelet basis  $\{\psi_\alpha^k\}$ ,  $k \in \mathbb{Z}$ ,  $y_\alpha^k \in \mathcal{Y}^k$ , of  $L^2(X)$ , having exponential decay*

$$(2.10) \quad |\psi_\alpha^k(x)| \leq \frac{C}{\sqrt{\mu(B(y_\alpha^k, \delta^k))}} \exp\left(-\nu\left(\frac{d(y_\alpha^k, x)}{\delta^k}\right)^a\right),$$

*Hölder regularity*

$$(2.11) \quad |\psi_\alpha^k(x) - \psi_\alpha^k(y)| \leq \frac{C}{\sqrt{\mu(B(y_\alpha^k, \delta^k))}} \left(\frac{d(x, y)}{\delta^k}\right)^\eta \exp\left(-\nu\left(\frac{d(y_\alpha^k, x)}{\delta^k}\right)^a\right)$$

*for  $d(x, y) \leq \delta^k$ , and the cancellation property*

$$(2.12) \quad \int_X \psi_\alpha^k(x) d\mu(x) = 0, \quad \text{for } k \in \mathbb{Z}, y_\alpha^k \in \mathcal{Y}^k.$$

Moreover, the wavelet expansion is given by

$$(2.13) \quad f(x) = \sum_{k \in \mathbb{Z}} \sum_{\alpha \in \mathcal{Y}^k} \langle f, \psi_\alpha^k \rangle \psi_\alpha^k(x)$$

in the sense of  $L^2(X)$ .

Here  $\delta$  is a fixed small parameter, say  $\delta \leq 10^{-3}A_0^{-10}$ , and  $C < \infty$ ,  $\nu > 0$  and  $\eta \in (0, 1]$  are constants independent of  $k$ ,  $\alpha$ ,  $x$  and  $y_\alpha^k$ .

In what follows, we refer to the functions  $\psi_\alpha^k$  as wavelets. Throughout this paper,  $a$  denotes the exponent from (2.9) and  $\eta$  denotes the Hölder-regularity exponent from (2.11).

**Remark 2.3.** The wavelets  $\{\psi_\alpha^k\}_{k, \alpha}$  form an unconditional basis of  $L^p(X)$  for  $1 < p < \infty$ , as shown in Corollary 10.4 in [AH]. Therefore, the reproducing formula (2.13) also holds for  $f \in L^p(X)$ . Moreover, for us the most crucial feature of the orthonormal wavelets construction is the following estimate, which is a special case of [AH, Lemma 8.3]:

$$(2.14) \quad \sum_{j \in \mathbb{Z}: \delta^j \geq r} \frac{1}{\mu(B(x, \delta^j))} \exp\left(-\nu\left(\frac{d(x, \mathcal{Y}^j)}{\delta^j}\right)^a\right) \leq \frac{C}{\mu(B(x, r))},$$

for all  $x \in X$  and  $r > 0$ , and for the constants  $\nu > 0$  and  $a := (1 + 2 \log_2 A_0)^{-1}$  from Theorem 2.2. Series of this type naturally arise in the context of proving that the reproducing formula holds for test functions and distributions, as well as in relation to function spaces. This estimate allows us to drop the extra assumption, used in previous work, of a reverse-doubling property on the measure.

Furthermore, this estimate is crucial for estimating the quantity  $\sum_k \sum_\alpha \psi_\alpha^k(x) \psi_\alpha^k(y)$ . We will use this estimate in the proofs of Theorems 3.4 and 4.3 below.

### 3. TEST FUNCTIONS, DISTRIBUTIONS, AND WAVELET REPRODUCING FORMULA

We now introduce test functions and distributions on spaces of homogeneous type  $(X, d, \mu)$  in the sense of Coifman and Weiss, and on product spaces  $(X_1, d_1, \mu_1) \times (X_2, d_2, \mu_2)$ . We show that the (scaled) Auscher–Hytönen wavelets are test functions (Theorem 3.3), and establish the wavelet reproducing formula for test functions and for distributions (Theorem 3.4, Corollary 3.5, Theorem 3.11). Along the way we establish some properties of the wavelet operators  $D_k$  (Lemma 3.6), and construct smooth cut-off functions using the splines of Auscher and Hytönen (Lemma 3.8).

We begin with the one-parameter case. For  $x, y \in X$  and  $r > 0$ , let

$$V_r(x) := \mu(B(x, r)) \quad \text{and} \quad V(x, y) := \mu(B(x, d(x, y))).$$

#### 3.1. One-parameter test functions, distributions, and wavelet reproducing formula.

**Definition 3.1.** (Test functions) Fix  $x_0 \in X$ ,  $r > 0$ ,  $\beta \in (0, \eta]$  where  $\eta$  is the regularity exponent from Theorem 2.2, and  $\gamma > 0$ . A function  $f$  defined on  $X$  is said to be a *test function of type  $(x_0, r, \beta, \gamma)$  centered at  $x_0 \in X$*  if  $f$  satisfies the following three conditions.

(i) (Size condition) For all  $x \in X$ ,

$$|f(x)| \leq C \frac{1}{V_r(x_0) + V(x, x_0)} \left( \frac{r}{r + d(x, x_0)} \right)^\gamma.$$

(ii) (Hölder regularity condition) For all  $x, y \in X$  with  $d(x, y) < (2A_0)^{-1}(r + d(x, x_0))$ ,

$$|f(x) - f(y)| \leq C \left( \frac{d(x, y)}{r + d(x, x_0)} \right)^\beta \frac{1}{V_r(x_0) + V(x, x_0)} \left( \frac{r}{r + d(x, x_0)} \right)^\gamma.$$

(iii) (Cancellation condition)

$$\int_X f(x) d\mu(x) = 0.$$

*A priori*, this definition makes sense for arbitrary  $\beta > 0$ . Here we have used the condition  $\beta \in (0, \eta]$  both for consistency with the earlier literature and since our focus is on the wavelets  $\psi_\alpha^k$ , which (when scaled) are test functions with  $\beta = \eta$ , as we will see.

We denote by  $G(x_0, r, \beta, \gamma)$  the set of all test functions of type  $(x_0, r, \beta, \gamma)$ . The norm of  $f$  in  $G(x_0, r, \beta, \gamma)$  is defined by

$$\|f\|_{G(x_0, r, \beta, \gamma)} := \inf\{C > 0 : \text{(i) and (ii) hold}\}.$$

For each fixed  $x_0$ , let  $G(\beta, \gamma) := G(x_0, 1, \beta, \gamma)$ . It is easy to check that for each fixed  $x'_0 \in X$  and  $r > 0$ , we have  $G(x'_0, r, \beta, \gamma) = G(\beta, \gamma)$  with equivalent norms. Furthermore, it is also easy to see that  $G(\beta, \gamma)$  is a Banach space with respect to the norm on  $G(\beta, \gamma)$ .

For  $\beta \in (0, \eta]$  and  $\gamma > 0$ , let  $\mathring{G}(\beta, \gamma)$  be the completion of the space  $G(\eta, \gamma)$  in the norm of  $G(\beta, \gamma)$ ; of course when  $\beta = \eta$  we simply have  $\mathring{G}(\beta, \gamma) = \mathring{G}(\eta, \gamma) = G(\eta, \gamma)$ . We define the norm on  $\mathring{G}(\beta, \gamma)$  by  $\|f\|_{\mathring{G}(\beta, \gamma)} := \|f\|_{G(\beta, \gamma)}$ .

It is immediate from the definition that the sets  $\mathring{G}(\beta, \gamma)$  are nested; for example if  $0 < \beta' < \beta$  and  $0 < \gamma' < \gamma$ , then  $\mathring{G}(\beta, \gamma) \subset \mathring{G}(\beta', \gamma')$ .

**Definition 3.2.** (Distributions) Fix  $x_0 \in X$ ,  $r > 0$ ,  $\beta \in (0, \eta]$  where  $\eta$  is the regularity exponent from Theorem 2.2, and  $\gamma > 0$ . The *distribution space*  $(\mathring{G}(\beta, \gamma))'$  is defined to be the set of all linear functionals  $\mathcal{L}$  from  $\mathring{G}(\beta, \gamma)$  to  $\mathbb{C}$  with the property that there exists  $C > 0$  such that for all  $f \in \mathring{G}(\beta, \gamma)$ ,

$$|\mathcal{L}(f)| \leq C \|f\|_{\mathring{G}(\beta, \gamma)}.$$

We note that for each  $\beta \in (0, \eta]$  and  $\gamma > 0$ , the set  $\mathring{G}(\beta, \gamma) \subset L^2(X)$ , while each  $f \in L^2(X)$  induces a distribution in  $(\mathring{G}(\beta, \gamma))'$ .

We now prove that the wavelets constructed in [AH], suitably scaled, are test functions.

**Theorem 3.3.** Suppose that  $\{\psi_\alpha^k\}_{k \in \mathbb{Z}, \alpha \in \mathcal{Y}^k}$  is an orthonormal wavelet basis as in Theorem 2.2 with Hölder regularity of order  $\eta$ . Then for each  $k \in \mathbb{Z}$ ,  $y_\alpha^k \in \mathcal{Y}^k$ , and  $\gamma > 0$ , the scaled wavelet  $\psi_\alpha^k(x)/\sqrt{\mu(B(y_\alpha^k, \delta^k))}$  belongs to the set  $G(y_\alpha^k, \delta^k, \eta, \gamma)$  of test functions of type  $(y_\alpha^k, \delta^k, \eta, \gamma)$  centered at  $y_\alpha^k$ .

Before proving Theorem 3.3, we make the following useful observation: by the doubling property (1.3) on the measure  $\mu$ , for each  $x_0, x \in X$  and  $r > 0$  with  $r \leq d(x_0, x)$ , we have  $V(x_0, x) \leq C(d(x_0, x)/r)^\omega V_r(x_0)$  and hence

$$(3.1) \quad \frac{1}{V_r(x_0)} \leq C \frac{1}{V_r(x_0) + V(x_0, x)} \left( \frac{d(x_0, x)}{r} \right)^\omega.$$

*Proof of Theorem 3.3.* By property (2.10) of  $\psi_\alpha^k$  from Theorem 2.2, we obtain that

$$\begin{aligned} \frac{\psi_\alpha^k(x)}{\sqrt{\mu(B(y_\alpha^k, \delta^k))}} &\leq \frac{C}{\mu(B(y_\alpha^k, \delta^k))} \exp(-\nu(\delta^{-k} d(y_\alpha^k, x))^a) \\ &\leq \frac{C}{V_{\delta^k}(y_\alpha^k)} \left( \frac{\delta^k}{\delta^k + d(y_\alpha^k, x)} \right)^\Gamma \end{aligned}$$

for all  $\Gamma > 0$ , with a constant  $C$  depending only on  $\nu$ ,  $a = (1 + 2 \log_2 A_0)^{-1}$ , and  $\Gamma$ . To see that  $\psi_\alpha^k(x)/\sqrt{\mu(B(y_\alpha^k, \delta^k))}$  satisfies the size condition Definition 3.1(i) for  $\gamma > 0$ , we consider two cases. First, if  $\delta^k > d(y_\alpha^k, x)$ , then  $V(y_\alpha^k, x) \leq V_{\delta^k}(y_\alpha^k)$  and hence

$$\frac{\psi_\alpha^k(x)}{\sqrt{\mu(B(y_\alpha^k, \delta^k))}} \leq \frac{C}{V_{\delta^k}(y_\alpha^k) + V(y_\alpha^k, x)} \left( \frac{\delta^k}{\delta^k + d(y_\alpha^k, x)} \right)^\Gamma.$$

For the second case, if  $\delta^k \leq d(y_\alpha^k, x)$ , an application of (3.1) shows that

$$\frac{\psi_\alpha^k(x)}{\sqrt{\mu(B(y_\alpha^k, \delta^k))}} \leq C \frac{1}{V_{\delta^k}(y_\alpha^k) + V(y_\alpha^k, x)} \left( \frac{\delta^k}{\delta^k + d(y_\alpha^k, x)} \right)^{\Gamma - \omega}.$$

Taking  $\Gamma > \omega$  and setting  $\gamma := \Gamma - \omega$ , we see that  $\psi_\alpha^k(x)/\sqrt{\mu(B(y_\alpha^k, \delta^k))}$  satisfies the size condition Definition 3.1(i) with  $x_0 = y_\alpha^k$  and  $r = \delta^k$ , and for arbitrary  $\gamma > 0$ .

We now show that  $\psi_\alpha^k(x)/\sqrt{\mu(B(y_\alpha^k, \delta^k))}$  satisfies the smoothness condition Definition 3.1(ii) for  $\gamma > 0$  and  $\beta = \eta$ , if  $d(x, y) \leq (2A_0)^{-1}(\delta^k + d(y_\alpha^k, x))$ . We consider three cases. First, suppose  $d(x, y) \leq \delta^k \leq (2A_0)^{-1}(\delta^k + d(y_\alpha^k, x))$ . Then property (2.11) yields

$$\left| \frac{\psi_\alpha^k(x) - \psi_\alpha^k(y)}{\sqrt{\mu(B(y_\alpha^k, \delta^k))}} \right| \leq C \left( \frac{d(x, y)}{\delta^k} \right)^\eta \frac{1}{\mu(B(y_\alpha^k, \delta^k))} \left( \frac{\delta^k}{\delta^k + d(y_\alpha^k, x)} \right)^\Gamma.$$

Note that in this case,  $\delta^k \leq (2A_0 - 1)^{-1}d(y_\alpha^k, x)$ , and so we may apply (3.1) with  $r = \delta^k$  and  $x_0 = y_\alpha^k$  to conclude that

$$\left| \frac{\psi_\alpha^k(x) - \psi_\alpha^k(y)}{\sqrt{\mu(B(y_\alpha^k, \delta^k))}} \right| \leq C \left( \frac{d(x, y)}{\delta^k + d(y_\alpha^k, x)} \right)^\eta \frac{1}{V_{\delta^k}(y_\alpha^k) + V(y_\alpha^k, x)} \left( \frac{\delta^k}{\delta^k + d(y_\alpha^k, x)} \right)^{\Gamma - \omega - \eta}.$$

Second, consider the case where  $\delta^k \leq d(x, y) \leq (2A_0)^{-1}(\delta^k + d(y_\alpha^k, x))$ . It is straightforward to verify from the quasi-triangle inequality that in this case,

$$(3.2) \quad \delta^k + d(y_\alpha^k, y) \sim \delta^k + d(y_\alpha^k, x).$$

Applying property (2.10) for  $\psi_\alpha^k(x)$  and  $\psi_\alpha^k(y)$ , we find

$$\begin{aligned} \left| \frac{\psi_\alpha^k(x) - \psi_\alpha^k(y)}{\sqrt{\mu(B(y_\alpha^k, \delta^k))}} \right| &\leq \frac{C}{V_{\delta^k}(y_\alpha^k)} \left( \frac{\delta^k}{\delta^k + d(y_\alpha^k, x)} \right)^\Gamma + \frac{C}{V_{\delta^k}(y_\alpha^k)} \left( \frac{\delta^k}{\delta^k + d(y_\alpha^k, y)} \right)^\Gamma \\ &\leq C \left( \frac{d(x, y)}{\delta^k} \right)^\eta \frac{1}{V_{\delta^k}(y_\alpha^k)} \left( \frac{\delta^k}{\delta^k + d(y_\alpha^k, x)} \right)^\Gamma \\ &\leq C \left( \frac{d(x, y)}{\delta^k + d(y_\alpha^k, x)} \right)^\eta \frac{1}{V_{\delta^k}(y_\alpha^k) + V(y_\alpha^k, x)} \left( \frac{\delta^k}{\delta^k + d(y_\alpha^k, x)} \right)^{\Gamma - \omega - \eta}. \end{aligned}$$

Here the second inequality follows from (3.2) and the fact that  $d(x, y)/\delta^k \geq 1$ , and the third inequality follows from (3.1).

For the third and last case, if  $\delta^k > (2A_0)^{-1}(\delta^k + d(y_\alpha^k, x)) \geq d(x, y)$ , then we have  $d(x, y) < \delta^k$  and  $d(y_\alpha^k, x) \leq (2A_0 - 1)\delta^k$ . Therefore, applying property (2.11) together with the fact that  $V(y_\alpha^k, x) \leq C\mu(B(y_\alpha^k, \delta^k))$  yields

$$\begin{aligned} \left| \frac{\psi_\alpha^k(x) - \psi_\alpha^k(y)}{\sqrt{\mu(B(y_\alpha^k, \delta^k))}} \right| &\leq C \left( \frac{d(x, y)}{\delta^k} \right)^\eta \frac{1}{V_{\delta^k}(y_\alpha^k)} \left( \frac{\delta^k}{\delta^k + d(y_\alpha^k, x)} \right)^\Gamma \\ &\leq C \left( \frac{d(x, y)}{\delta^k + d(y_\alpha^k, x)} \right)^\eta \frac{1}{V_{\delta^k}(y_\alpha^k) + V(y_\alpha^k, x)} \left( \frac{\delta^k}{\delta^k + d(y_\alpha^k, x)} \right)^{\Gamma - \eta}. \end{aligned}$$

Combining all the cases above, in the first and second cases take  $\Gamma > \omega + \eta$  and set  $\gamma := \Gamma - \omega - \eta$ , and in the third case take  $\Gamma > \eta$  and set  $\gamma := \Gamma - \eta$ . We see that the function  $\psi_\alpha^k(x)/\sqrt{\mu(B(y_\alpha^k, \delta^k))}$  satisfies the smoothness condition (ii) in Definition 3.1 with  $x_0 = y_\alpha^k$ ,  $r = \delta^k$ ,  $\beta = \eta$  and for arbitrary  $\gamma > 0$ .

Moreover, the cancellation property of the function  $\psi_\alpha^k(x)/\sqrt{\mu(B(y_\alpha^k, \delta^k))}$  is immediate from property (2.12) of  $\psi_\alpha^k$ , by Theorem 2.2. Thus,  $\psi_\alpha^k(x)/\sqrt{\mu(B(y_\alpha^k, \delta^k))}$  belongs to the test function space  $G(y_\alpha^k, \delta^k, \eta, \gamma)$ . This completes the proof of Theorem 3.3.  $\square$

Now we state and prove the main result of this subsection, which will be the crucial tool for establishing the Littlewood–Paley theory and developing the Hardy spaces.

**Theorem 3.4.** (Wavelet reproducing formula for test functions) *Suppose that  $f \in \mathring{G}(\beta, \gamma)$  with  $\beta, \gamma \in (0, \eta)$ . Then the wavelet reproducing formula*

$$(3.3) \quad f(x) = \sum_{k \in \mathbb{Z}} \sum_{\alpha \in \mathcal{Y}^k} \langle f, \psi_\alpha^k \rangle \psi_\alpha^k(x)$$

*holds in  $\mathring{G}(\beta', \gamma')$  for all  $\beta' \in (0, \beta)$  and  $\gamma' \in (0, \gamma)$ .*

As an immediate consequence of Theorem 3.4, we obtain the following result.

**Corollary 3.5.** (Wavelet reproducing formula for distributions) *Take  $\beta, \gamma \in (0, \eta)$ . Then the wavelet reproducing formula (3.3) also holds in the space  $(\mathring{G}(\beta, \gamma))'$  of distributions.*

*Proof of Theorem 3.4.* Take  $f \in \mathring{G}(\beta, \gamma)$  with  $\beta, \gamma \in (0, \eta)$ . It suffices to show that

$$(3.4) \quad \left\| \sum_{|k| > L} \sum_{\alpha \in \mathcal{Y}^k} \langle \psi_\alpha^k, f \rangle \psi_\alpha^k(\cdot) \right\|_{G(\beta', \gamma')} \longrightarrow 0$$

as  $L$  tends to infinity, for each  $\beta' \in (0, \beta)$  and  $\gamma' \in (0, \gamma)$ .

The proof of (3.4) is based on the following estimate: for each  $\beta' \in (0, \beta)$  and  $\gamma' \in (0, \gamma)$ , there is a constant  $\sigma > 0$  such that for each  $L \in \mathbb{N}$ ,

$$(3.5) \quad \left\| \sum_{|k| > L} \sum_{\alpha \in \mathcal{Y}^k} \langle \psi_\alpha^k, f \rangle \psi_\alpha^k(\cdot) \right\|_{G(\beta', \gamma')} \leq C \delta^{\sigma L} \|f\|_{G(\beta, \gamma)},$$

where  $C$  is a positive constant independent of  $f \in \mathring{G}(\beta, \gamma)$ .

To verify (3.5), it suffices to show that the following decay and smoothness estimates hold: for each  $\gamma' \in (0, \gamma)$ , there exist a positive constant  $C$  independent of  $f$ , and a positive number  $\sigma'$ , such that

$$(3.6) \quad \left| \sum_{|k| > L} \sum_{\alpha \in \mathcal{Y}^k} \langle \psi_\alpha^k, f \rangle \psi_\alpha^k(x) \right| \leq C \delta^{\sigma' L} \frac{1}{V_1(x_0) + V(x, x_0)} \left( \frac{1}{1 + d(x, x_0)} \right)^{\gamma'} \|f\|_{G(\beta, \gamma)}, \quad \text{and}$$

$$(3.7) \quad \left| \sum_{|k| > L} \sum_{\alpha \in \mathcal{Y}^k} \langle \psi_\alpha^k, f \rangle \psi_\alpha^k(x) - \sum_{|k| > L} \sum_{\alpha \in \mathcal{Y}^k} \langle \psi_\alpha^k, f \rangle \psi_\alpha^k(x') \right| \\ \leq C \left( \frac{d(x, x')}{1 + d(x, x_0)} \right)^\beta \frac{1}{V_1(x_0) + V(x, x_0)} \left( \frac{1}{1 + d(x, x_0)} \right)^\gamma \|f\|_{G(\beta, \gamma)}$$

for all  $x$  and  $x'$  such that  $d(x, x') \leq (2A_0)^{-1}(1 + d(x, x_0))$ .

Indeed, to see that (3.6) and (3.7) imply (3.5), we take the geometric mean between (3.7) and the following estimate (obtained directly from (3.6)):

$$\left| \sum_{|k| > L} \sum_{\alpha \in \mathcal{Y}^k} \langle \psi_\alpha^k, f \rangle \psi_\alpha^k(x) - \sum_{|k| > L} \sum_{\alpha \in \mathcal{Y}^k} \langle \psi_\alpha^k, f \rangle \psi_\alpha^k(x') \right| \\ \leq \left| \sum_{|k| > L} \sum_{\alpha \in \mathcal{Y}^k} \langle \psi_\alpha^k, f \rangle \psi_\alpha^k(x) \right| + \left| \sum_{|k| > L} \sum_{\alpha \in \mathcal{Y}^k} \langle \psi_\alpha^k, f \rangle \psi_\alpha^k(x') \right| \\ \leq C \delta^{\sigma' L} \frac{1}{V_1(x_0) + V(x, x_0)} \left( \frac{1}{1 + d(x, x_0)} \right)^{\gamma'} \|f\|_{G(\beta, \gamma)}.$$

This gives

$$(3.8) \quad \left| \sum_{|k| > L} \sum_{\alpha \in \mathcal{Y}^k} \langle \psi_\alpha^k, f \rangle \psi_\alpha^k(x) - \sum_{|k| > L} \sum_{\alpha \in \mathcal{Y}^k} \langle \psi_\alpha^k, f \rangle \psi_\alpha^k(x') \right|$$

$$\leq C\delta^{\sigma L} \left( \frac{d(x, x')}{r + d(x, x_0)} \right)^{\beta'} \frac{1}{V_1(x_0) + V(x, x_0)} \left( \frac{1}{1 + d(x, x_0)} \right)^{\gamma'} \|f\|_{G(\beta, \gamma)}$$

for some  $\sigma < \sigma'$ . Now (3.6) and (3.8), together with the fact that

$$\int_X \sum_{|k| > L} \sum_{\alpha \in \mathcal{Y}^k} \langle \psi_\alpha^k, f \rangle \psi_\alpha^k(x) d\mu(x) = 0,$$

imply that  $\sum_{|k| > L} \sum_{\alpha \in \mathcal{Y}^k} \langle \psi_\alpha^k, f \rangle \psi_\alpha^k$  is a test function in  $G(\beta', \gamma')$ . Moreover, we see from the upper bounds in (3.6) and (3.8) that (3.5) holds, as required.

To prove the decay and smoothness estimates (3.6) and (3.7), we need the following lemma which gives estimates for the kernels  $D_k(x, y) = \sum_{\alpha \in \mathcal{Y}^k} \psi_\alpha^k(x) \psi_\alpha^k(y)$  of the wavelet operators  $D_k$ . These wavelet operators are defined by

$$D_k f(x) := \sum_{\alpha \in \mathcal{Y}^k} \langle \psi_\alpha^k, f \rangle \psi_\alpha^k(x) = \int_X D_k(x, y) f(y) dy,$$

for  $k \in \mathbb{Z}$ . We note that the first two estimates in Lemma 3.6 are similar to estimates given in Lemma 9.1 in [AH].

**Lemma 3.6.** (Properties of wavelet operators  $D_k$ ) *Let*

$$D_k(x, y) := \sum_{\alpha \in \mathcal{Y}^k} \psi_\alpha^k(x) \psi_\alpha^k(y)$$

*for  $x, y \in X$ . Fix  $\gamma > 0$ . Then the following estimates hold.*

(i) (Decay condition) *For all  $x, y \in X$ , we have*

$$(3.9) \quad |D_k(x, y)| \leq C \frac{1}{V_{\delta^k}(x) + V(x, y)} \left( \frac{\delta^k}{\delta^k + d(x, y)} \right)^\gamma.$$

(ii) (Smoothness condition) *If  $d(y, y') \leq (2A_0)^{-1} \max\{\delta^k + d(x, y), \delta^k + d(x, y')\}$ , then*

$$(3.10) \quad |D_k(x, y) - D_k(x, y')| \leq C \left( \frac{d(y, y')}{\delta^k + d(x, y)} \right)^\eta \frac{1}{V_{\delta^k}(x) + V(x, y)} \left( \frac{\delta^k}{\delta^k + d(x, y)} \right)^\gamma.$$

*The same estimate holds with  $x$  and  $y$  interchanged.*

(iii) (Double smoothness condition) *If*

$$\begin{aligned} d(x, x') &\leq (2A_0)^{-1} \max\{\delta^k + d(x, y), \delta^k + d(x', y)\} \quad \text{and} \\ d(y, y') &\leq (2A_0)^{-1} \max\{\delta^k + d(x, y), \delta^k + d(x, y')\}, \end{aligned}$$

*then*

$$(3.11) \quad \begin{aligned} &|D_k(x, y) - D_k(x', y) - D_k(x, y') + D_k(x', y')| \\ &\leq C \left( \frac{d(x, x')}{\delta^k + d(x, y)} \right)^\eta \left( \frac{d(y, y')}{\delta^k + d(x, y)} \right)^\eta \frac{1}{V_{\delta^k}(x) + V(x, y)} \left( \frac{\delta^k}{\delta^k + d(x, y)} \right)^\gamma. \end{aligned}$$

We defer the proof of Lemma 3.6 until after the end of the proof of Theorem 3.4.

Returning to the proof of Theorem 3.4, we first show (3.6). Write

$$\begin{aligned} \left| \sum_{|k| > L} \sum_{\alpha \in \mathcal{Y}^k} \langle \psi_\alpha^k, f \rangle \psi_\alpha^k(x) \right| &\leq \left| \sum_{k > L} \sum_{\alpha \in \mathcal{Y}^k} \langle \psi_\alpha^k, f \rangle \psi_\alpha^k(x) \right| + \left| \sum_{k < -L} \sum_{\alpha \in \mathcal{Y}^k} \langle \psi_\alpha^k, f \rangle \psi_\alpha^k(x) \right| \\ &=: (A) + (B). \end{aligned}$$

For (A), using the cancellation property (2.12) of the wavelet  $\psi_\alpha^k$  and integrating over the sets  $W_1 := \{y \in X : d(x, y) \leq (2A_0)^{-1}(1 + d(x, x_0))\}$  and  $W_2 := X \setminus W_1$ , we obtain

$$\begin{aligned} (A) &\leq \sum_{k>L} \int_{W_1} |D_k(x, y)| |f(y) - f(x)| d\mu(y) \\ &\quad + \sum_{k>L} \int_{W_2} |D_k(x, y)| (|f(y)| + |f(x)|) d\mu(y) \\ &=: (A)_1 + (A)_2. \end{aligned}$$

To deal with  $(A)_1$ , applying the decay condition (3.9) from Lemma 3.6 on  $D_k(x, y)$  and the Hölder regularity property (Definition 3.1(ii)) of the test function  $f \in G(\beta, \gamma)$  gives

$$\begin{aligned} (A)_1 &\leq C \|f\|_{G(\beta, \gamma)} \sum_{k>L} \int_{W_1} \frac{1}{V_{\delta^k}(x) + V(x, y)} \left( \frac{\delta^k}{\delta^k + d(x, y)} \right)^\gamma \\ &\quad \times \left( \frac{d(x, y)}{1 + d(x, x_0)} \right)^\beta \frac{1}{V_1(x_0) + V(x, x_0)} \left( \frac{1}{1 + d(x, x_0)} \right)^\gamma d\mu(y) \\ &\leq C \delta^{\beta L} \frac{1}{V_1(x_0) + V(x, x_0)} \left( \frac{1}{1 + d(x, x_0)} \right)^\gamma \|f\|_{G(\beta, \gamma)}. \end{aligned}$$

To estimate  $(A)_2$ , applying the size conditions on both  $D_k(x, y)$  (Lemma 3.6) and  $f$  (Definition 3.1(i)) gives

$$\begin{aligned} (A)_2 &\leq C \|f\|_{G(\beta, \gamma)} \sum_{k>L} \int_{W_2} \frac{1}{V_{\delta^k}(x) + V(x, y)} \left( \frac{\delta^k}{\delta^k + d(x, y)} \right)^\gamma \\ &\quad \times \left[ \frac{1}{V_1(x_0) + V(y, x_0)} \left( \frac{1}{1 + d(y, x_0)} \right)^\gamma + \frac{1}{V_1(x_0) + V(x, x_0)} \left( \frac{1}{1 + d(x, x_0)} \right)^\gamma \right] d\mu(y). \end{aligned}$$

For the first sum, use the fact that if  $d(x, y) > (2A_0)^{-1}(1 + d(x, x_0))$  then  $V(x, y) \geq CV(x, (2A_0)^{-1}(1 + d(x, x_0))) \geq C[V_1(x_0) + V(x, x_0)]$ . For the second sum, apply the following estimate:

$$\int_{W_2} \frac{1}{V_{\delta^k}(x) + V(x, y)} \left( \frac{\delta^k}{\delta^k + d(x, y)} \right)^\gamma \mu(y) \leq C \delta^{\gamma k}.$$

We obtain

$$(A)_2 \leq C \|f\|_{G(\beta, \gamma)} \delta^{\gamma L} \frac{1}{V_1(x_0) + V(x, x_0)} \left( \frac{1}{1 + d(x, x_0)} \right)^\gamma.$$

Hence for some  $\sigma > 0$ ,

$$(A) \leq C \delta^{\sigma L} \frac{1}{V_1(x_0) + V(x, x_0)} \left( \frac{1}{1 + d(x, x_0)} \right)^\gamma \|f\|_{G(\beta, \gamma)}.$$

We now turn to (B). Using the fact that  $\int_X f(x) d\mu(x) = 0$  and considering the sets  $W_3 := \{y \in X : d(y, x_0) \leq (2A_0)^{-1}(\delta^k + d(x, x_0))\}$  and  $W_4 := X \setminus W_3$ , we have

$$\begin{aligned} (B) &\leq \sum_{k<-L} \int_{W_3} |D_k(x, y) - D_k(x, x_0)| |f(y)| d\mu(y) \\ &\quad + \sum_{k<-L} \int_{W_4} (|D_k(x, y)| + |D_k(x, x_0)|) |f(y)| d\mu(y) \\ &=: (B)_1 + (B)_2. \end{aligned}$$

For  $(B)_1$ , applying the smoothness estimate from Lemma 3.6(ii) and the size estimate of the test function  $f$  (Definition 3.1(i)) yields

$$\begin{aligned} (B)_1 &\leq C \|f\|_{G(\beta, \gamma)} \sum_{k < -L} \int_{W_3} \left( \frac{d(y, x_0)}{\delta^k + d(x, x_0)} \right)^{\eta'} \frac{1}{V_{\delta^k}(x) + V(x, x_0)} \\ &\quad \times \left( \frac{\delta^k}{\delta^k + d(x, x_0)} \right)^{\gamma'} \frac{1}{V_1(x_0) + V(y, x_0)} \left( \frac{1}{1 + d(y, x_0)} \right)^{\gamma} d\mu(y) \\ &\leq C \|f\|_{G(\beta, \gamma)} \delta^{(\eta' - \gamma')L} \frac{1}{V_1(x_0) + V(x, x_0)} \left( \frac{1}{1 + d(x, x_0)} \right)^{\gamma'}, \end{aligned}$$

where we choose  $\gamma' < \eta' < \gamma$ . To estimate  $(B)_2$ , we first write  $(B)_2 = (B)_{21} + (B)_{22}$  where

$$(B)_{21} := \sum_{k < -L} \int_{W_4} |D_k(x, y)| |f(y)| d\mu(y)$$

and

$$(B)_{22} := \sum_{k < -L} \int_{W_4} |D_k(x, x_0)| |f(y)| d\mu(y).$$

Since here  $d(y, x_0) > (2A_0)^{-1}(\delta^k + d(x, x_0))$ , the size estimates for the test function  $f$  (Definition 3.1(i)) imply that for  $0 < \gamma' < \gamma$ ,

$$\begin{aligned} |f(y)| &\leq C \|f\|_{G(\beta, \gamma)} \frac{1}{V_1(x_0) + V(y, x_0)} \left( \frac{1}{1 + d(y, x_0)} \right)^{\gamma} \\ &\leq C \|f\|_{G(\beta, \gamma)} \delta^{k(\gamma' - \gamma)} \frac{1}{V_1(x_0) + V(x, x_0)} \left( \frac{1}{1 + d(x, x_0)} \right)^{\gamma'}. \end{aligned}$$

The above estimate, together with the fact that  $\int_X |D_k(x, y)| d\mu(y) \leq C$ , yields

$$(B)_{21} \leq C \delta^{(\gamma - \gamma')L} \frac{1}{V_1(x_0) + V(x, x_0)} \left( \frac{1}{1 + d(x, x_0)} \right)^{\gamma'} \|f\|_{G(\beta, \gamma)}.$$

The estimate for  $(B)_{22}$  is similar, but easier. Indeed, by Lemma 3.6, we have

$$|D_k(x, x_0)| \leq C \frac{1}{V_{\delta^k}^k(x_0) + V(x, x_0)} \left( \frac{\delta^k}{\delta^k + d(x, x_0)} \right)^{\gamma'}$$

and

$$\int_{W_4} |f(y)| d\mu(y) \leq C \left( \frac{1}{\delta^k + d(x, x_0)} \right)^{\gamma}.$$

Thus, we obtain the same estimate for  $(B)_{22}$  as for  $(B)_{21}$ , but with  $\gamma - \gamma'$  replaced by  $\eta' - \gamma'$ . This completes the proof of (3.6).

Finally, we show (3.7). To do this, we first need to construct a smooth cut-off function. For this purpose, we recall the following result on the properties of the *spline functions*  $s_{\alpha}^k$  on  $(X, d, \mu)$  that were constructed by Auscher and Hytönen [AH].

**Theorem 3.7** ([AH], Theorem 3.1). *The spline functions  $s_{\alpha}^k$  satisfy the following properties: bounded support*

$$\chi_{B(x_{\alpha}^k, 1/8A_0^{-3}\delta^k)}(x) \leq s_{\alpha}^k(x) \leq \chi_{B(x_{\alpha}^k, 8A_0^5\delta^k)}(x);$$

*the interpolation and reproducing properties*

$$s_{\alpha}^k(x_{\beta}^k) = \delta_{\alpha, \beta}, \quad \sum_{\alpha} s_{\alpha}^k(x) = 1, \quad s_{\alpha}^k = \sum_{\beta} p_{\alpha, \beta}^k s_{\beta}^{k+1}(x),$$



where  $\{p_{\alpha,\beta}^k\}_\beta$  is a finite nonzero set of nonnegative coefficients with  $p_{\alpha,\beta}^k \leq 1$ ; and Hölder continuity

$$|s_\alpha^k(x) - s_\alpha^k(y)| \leq C \left( \frac{d(x,y)}{\delta^k} \right)^\eta.$$

We point out that in the above theorem,  $\alpha$  runs over  $\mathcal{X}^k$ . Using these splines we can construct a smooth cut-off function, as follows.

**Lemma 3.8.** (Smooth cut-off function) *For each fixed  $x_0 \in X$  and  $R_0 \in (0, \infty)$ , there exists a smooth cut-off function  $h(x)$  such that  $0 \leq h(x) \leq 1$ ,*

$$h(x) \equiv 1 \quad \text{when } x \in B(x_0, R_0/4), \quad h(x) \equiv 0 \quad \text{when } x \in B(x_0, A_0^2 R_0)^c,$$

and there exists a positive constant  $C$  independent of  $x_0, R_0, x, y$  such that

$$|h(x) - h(y)| \leq C \left( \frac{d(x,y)}{\delta^k} \right)^\eta.$$

*Proof.* For a fixed  $R_0 \in (0, \infty)$ , we choose  $k_0 \in \mathbb{Z}$  such that

$$8A_0^5 \delta^{k_0} \leq R_0/4 \quad \text{and} \quad 8A_0^5 \delta^{k_0-1} > R_0/4.$$

Next, we define the index set  $\mathcal{I}_{k_0}$  as follows:

$$\mathcal{I}_{k_0} := \{ \alpha \in \mathcal{X}^{k_0} : B(x_\alpha^{k_0}, 8A_0^5 \delta^{k_0}) \cap B(x_0, R_0/4) \neq \emptyset \}.$$

Then the number of indices contained in  $\mathcal{I}_{k_0}$  is bounded by a constant independent of  $R_0, k_0$ , and  $x_0$ , since  $8A_0^5 \delta^{k_0}$  is comparable to  $R_0$  and the reference dyadic points  $\{x_\alpha^{k_0}\}$  are  $\delta^{k_0}$ -separated.

Now define

$$h(x) := \sum_{\alpha \in \mathcal{I}_{k_0}} s_\alpha^{k_0}(x).$$

From the properties of the spline functions  $s_\alpha^k(x)$  (Theorem 3.7), it is easy to verify that  $h(x)$  satisfies all the properties listed in Lemma 3.8.  $\square$

In what follows, the Hölder-regularity index of the cut-off function  $h(x)$  is the  $\eta$  given in Theorem 3.7 ([AH], Theorem 3.1).

We now return to the proof of (3.7). It suffices to show that there exists a constant  $C$  such that for each  $M > 0$  and for  $d(x, x') \leq (2A_0)^{-1}(1 + d(x, x_0))$ ,

$$(3.12) \quad \left| \int_X \sum_{|k| \leq M} [D_k(x, y) - D_k(x', y)] f(y) d\mu(y) \right| \\ \leq C \left( \frac{d(x, x')}{1 + d(x, x_0)} \right)^\beta \frac{1}{V_1(x_0) + V(x, x_0)} \left( \frac{1}{1 + d(x, x_0)} \right)^\gamma.$$

To do this, fix  $M > 0$  and let  $T$  denote the wavelet operator given by

$$T(f)(x) := \int_X K(x, y) f(y) d\mu(y),$$

with kernel  $K(x, y) := \sum_{|k| \leq M} D_k(x, y)$ , where the kernel  $D_k(x, y)$  of the wavelet operator  $D_k$  is given by  $\sum_{\alpha \in \mathcal{X}^k} \psi_\alpha^k(x) \psi_\alpha^k(y)$ .

To show (3.12), it suffices to prove that if  $d(x, x') \leq (2A_0)^{-1}(1 + d(x, x_0))$ , then

$$|T(f)(x) - T(f)(x')| \leq C \left( \frac{d(x, x')}{1 + d(x, x_0)} \right)^\beta \frac{1}{V_1(x_0) + V(x, x_0)} \left( \frac{1}{1 + d(x, x_0)} \right)^\gamma.$$

Let  $R = d(x, x_0)$  and  $r = d(x, x')$ , and consider the case where  $R \geq 10$  and  $r \leq (20A_0^2)^{-1}(1 + d(x, x_0))$ . Following [M2], set  $1 = I(y) + J(y) + L(y)$ , where  $I(y)$  is a smooth cut-off function as in Lemma 3.8, satisfying

$$I(y) \equiv 1 \quad \text{when } y \in B(x, R/32A_0^2) \quad \text{and} \quad I(y) \equiv 0 \quad \text{when } y \in B(x, R/8)^c,$$

and

$$J(y) \equiv 1 \quad \text{when } y \in B(x_0, R/32A_0^2) \quad \text{and} \quad I(y) \equiv 0 \quad \text{when } y \in B(x_0, R/8)^c.$$

Also set

$$f_1(y) := f(y)I(y), \quad f_2(y) := f(y)J(y) \quad \text{and} \quad f_3(y) := f(y)L(y).$$

It is easy to see that  $f_1, f_2$  and  $f_3$  satisfy the following estimates (3.13)–(3.17):

$$(3.13) \quad |f_1(y)| \leq C \|f\|_{G(\beta, \gamma)} \frac{1}{V_1(x_0) + V(x, x_0)} \left( \frac{1}{1 + d(x, x_0)} \right)^\gamma$$

since  $|f_1(y)| \leq |f(y)|$  and  $1 + d(y, x_0) \geq C(1 + d(x, x_0))$  by the form of  $f_1$ ;

$$(3.14) \quad \begin{aligned} |f_1(y) - f_1(y')| &\leq |I(y)| |f(y) - f(y')| + |f(y')| |I(y) - I(y')| \\ &\leq C \|f\|_{G(\beta, \gamma)} \left( \frac{d(y, y')}{1 + d(x, x_0)} \right)^\beta \frac{1}{V_1(x_0) + V(x, x_0)} \left( \frac{1}{1 + d(x, x_0)} \right)^\gamma \end{aligned}$$

for all  $y$  and  $y'$ ;

$$(3.15) \quad |f_3(y)| \leq C \|f\|_{G(\beta, \gamma)} \frac{1}{V_1(x_0) + V(y, x_0)} \left( \frac{1}{1 + d(y, x_0)} \right)^\gamma \chi_{\{y \in X : d(y, x_0) > R/8\}};$$

$$(3.16) \quad \int_X |f_3(y)| d\mu(y) \leq C \|f\|_{G(\beta, \gamma)} \left( \frac{1}{1 + d(x, x_0)} \right)^\gamma;$$

and

$$(3.17) \quad \begin{aligned} \left| \int_X f_2(y) d\mu(y) \right| &= \left| - \int_X f_1(y) d\mu(y) - \int_X f_3(y) d\mu(y) \right| \\ &\leq C \|f\|_{G(\beta, \gamma)} \left( \frac{1}{1 + d(x, x_0)} \right)^\gamma. \end{aligned}$$

We write

$$\begin{aligned} T(f_1)(x) &= \underbrace{\int_X K(x, y) u(y) [f_1(y) - f_1(x)] d\mu(y)}_{p(x)} \\ &\quad + \underbrace{\int_X K(x, y) v(y) f_1(y) d\mu(y) + f_1(x) \int_X K(x, y) u(y) d\mu(y)}_{q(x)}, \end{aligned}$$

where  $u(y)$  is a smooth cut-off function as in Lemma 3.8, satisfying  $u(y) \equiv 1$  when  $y \in B(x, r)$ , and  $u(y) \equiv 0$  when  $y \in B(x, 4A_0^2 r)^c$ , and where  $v(y) := 1 - u(y)$ .

In order to estimate the expressions  $p(x)$  and  $q(x)$ , we show that the kernel  $K(x, y)$  of  $T$  satisfies the following four estimates:

$$(3.18) \quad |K(x, y)| \leq \frac{C}{V(x, y)}$$

for all  $x \neq y$ ;

$$(3.19) \quad |K(x, y) - K(x', y)| \leq \frac{C}{V(x, y)} \left( \frac{d(x, x')}{d(x, y)} \right)^\eta$$

for  $d(x, x') \leq (2A_0)^{-1}d(x, y)$ ;

$$(3.20) \quad |K(x, y) - K(x, y')| \leq \frac{C}{V(x, y)} \left( \frac{d(y, y')}{d(x, y)} \right)^\eta$$

for  $d(y, y') \leq (2A_0)^{-1}d(x, y)$ ; and

$$(3.21) \quad |K(x, y) - K(x, y') - K(x, y') + K(x', y')| \leq \frac{C}{V(x, y)} \left( \frac{d(x, x')}{d(x, y)} \right)^\eta \left( \frac{d(y, y')}{d(x, y)} \right)^\eta$$

for  $d(x, x') \leq (2A_0)^{-1}d(x, y)$  and  $d(y, y') \leq (2A_0)^{-1}d(x, y)$ .

We point out that without additional assumptions on the measure  $\mu$ , the decay and smoothness estimates (3.9)–(3.10) for  $D_k(x, y)$  as given in Lemma 3.6 are not by themselves sufficient to imply the above estimates for  $K(x, y) = \sum_k D_k(x, y)$ . Fortunately, however, in our setting, although we have no additional assumptions on  $\mu$ , we do have the special form  $K(x, y) = \sum_k D_k(x, y) = \sum_{k, \alpha} \psi_\alpha^k(x) \psi_\alpha^k(y)$  in terms of the wavelet basis, rather than the operators  $D_k = S_{k+1} - S_k$  from the classical case. Instead of using the estimates (3.9)–(3.10) directly, we can use the estimate (3.22) proved below, together with the approach of Lemma 9.1 in [AH]. The estimates (3.18)–(3.20) for our  $K(x, y)$  were proved in Lemmas 9.2 and 9.3 of [AH]. Thus, we only need to show the estimate (3.21). To do this, following the approach of Lemma 9.3 in [AH], we first claim that if  $d(x, x') \leq (2A_0)^{-1}d(x, y)$  and  $d(y, y') \leq (2A_0)^{-1}d(x, y)$ , then

$$(3.22) \quad \sum_{\alpha \in \mathcal{Y}^k} \left| [\psi_\alpha^k(x) - \psi_\alpha^k(x')] [\psi_\alpha^k(y) - \psi_\alpha^k(y')] \right| \\ \leq \frac{C}{V(x, \delta^k)} \min \left\{ 1, \left( \frac{d(x, x')}{\delta^k} \right)^\eta \right\} \min \left\{ 1, \left( \frac{d(y, y')}{\delta^k} \right)^\eta \right\} \\ \times \exp(-\nu(\delta^{-k}d(x, \mathcal{Y}^k))^a) \exp(-\nu(\delta^{-k}d(x, y))^a).$$

Recall that  $a := (1 + \log_2 A_0)^{-1}$  is the exponent defined in (2.9) in Theorem 2.2.

We prove (3.22) following the method used to prove the second assertion in Lemma 9.1 of [AH]. We consider four cases. First suppose  $\delta^k \geq d(x, x')$  and  $\delta^k \geq d(y, y')$ . Then

$$\left| [\psi_\alpha^k(x) - \psi_\alpha^k(x')] [\psi_\alpha^k(y) - \psi_\alpha^k(y')] \right| \\ \leq \frac{C}{V(x, \delta^k)} \left( \frac{d(x, x')}{\delta^k} \right)^\eta \left( \frac{d(y, y')}{\delta^k} \right)^\eta \exp(-\nu(\delta^{-k}d(x, y_\alpha^k))^a) \exp(-\nu(\delta^{-k}d(y, y_\alpha^k))^a) \\ \leq \frac{C}{V(x, \delta^k)} \left( \frac{d(x, x')}{\delta^k} \right)^\eta \left( \frac{d(y, y')}{\delta^k} \right)^\eta \exp(-\nu(\delta^{-k}d(x, y_\alpha^k))^a) \exp(-\nu(\delta^{-k}d(x, y))^a).$$

(Here the constant  $\nu$  changes from line to line to accommodate the constant  $A_0$  that arises from the use of the quasi-triangle inequality.) The sum over  $\alpha \in \mathcal{Y}^k$  of the fourth factor on the right-hand side is dominated by the expression  $\exp(-\nu(\delta^{-k}d(x, \mathcal{Y}^k))^a)$ .

Second, suppose  $\delta^k \geq d(x, x')$  and  $\delta^k < d(y, y') \leq (2A_0)^{-1}d(x, y)$ . Then we have

$$\begin{aligned} & \left| [\psi_\alpha^k(x) - \psi_\alpha^k(x')] [\psi_\alpha^k(y) - \psi_\alpha^k(y')] \right| \\ & \leq \left| [\psi_\alpha^k(x) - \psi_\alpha^k(x')] \psi_\alpha^k(y) \right| + \left| [\psi_\alpha^k(x) - \psi_\alpha^k(x')] \psi_\alpha^k(y') \right|. \end{aligned}$$

Then, by using the second estimate in Lemma 9.1 from [AH], and the quasi-triangle inequality, we obtain that

$$\begin{aligned} & \left| [\psi_\alpha^k(x) - \psi_\alpha^k(x')] [\psi_\alpha^k(y) - \psi_\alpha^k(y')] \right| \\ & \leq \frac{C}{V(x, \delta^k)} \left( \frac{d(x, x')}{\delta^k} \right)^\eta \exp(-\nu(\delta^{-k}d(x, y_\alpha^k))^a) \exp(-\nu(\delta^{-k}d(x, y))^a). \end{aligned}$$

Here the sum over  $\alpha \in \mathcal{Y}^k$  of the third factor on the right-hand side is dominated by  $\exp(-\nu(\delta^{-k}d(x, \mathcal{Y}^k))^a)$ .

The other two cases, namely when  $\delta^k < d(x, x') \leq (2A_0)^{-1}d(x, y)$  and  $\delta^k \geq d(y, y')$ , and when  $\delta^k < d(x, x') \leq (2A_0)^{-1}d(x, y)$  and  $\delta^k < d(y, y') \leq (2A_0)^{-1}d(x, y)$ , can be handled similarly. We omit the details.

Combining the estimates for all four cases above, we have established the claim (3.22).

Now we verify (3.21). From the definition of  $K(x, y)$  and the claim (3.22), we see that

$$\begin{aligned} & |K(x, y) - K(x, y') - K(x, y') + K(x', y')| \\ & \leq \sum_k \sum_{\alpha \in \mathcal{Y}^k} \left| [\psi_\alpha^k(x) - \psi_\alpha^k(x')] [\psi_\alpha^k(y) - \psi_\alpha^k(y')] \right| \\ & \leq \sum_{k: \delta^k \geq (2A_0)^{-1}d(x, y)} \frac{C}{V(x, \delta^k)} \left( \frac{d(x, x')}{\delta^k} \right)^\eta \left( \frac{d(y, y')}{\delta^k} \right)^\eta \exp(-\nu(\delta^{-k}d(x, \mathcal{Y}^k))^a) \\ & \quad + \sum_{\substack{k: \delta^k < (2A_0)^{-1}d(x, y), \\ d(x, x') \leq \delta^k, d(y, y') \leq \delta^k}} \frac{C}{V(x, y)} \left( \frac{d(x, y)}{\delta^k} \right)^\omega \left( \frac{d(x, x')}{\delta^k} \right)^\eta \left( \frac{d(y, y')}{\delta^k} \right)^\eta \exp(-\nu(\delta^{-k}d(x, y))^a) \\ & \quad + \sum_{\substack{k: \delta^k < (2A_0)^{-1}d(x, y), \\ d(x, x') > \delta^k, d(y, y') \leq \delta^k}} \frac{C}{V(x, y)} \left( \frac{d(x, y)}{\delta^k} \right)^\omega \left( \frac{d(y, y')}{\delta^k} \right)^\eta \exp(-\nu(\delta^{-k}d(x, y))^a) \\ & \quad + \sum_{\substack{k: \delta^k < (2A_0)^{-1}d(x, y), \\ d(x, x') \leq \delta^k, d(y, y') > \delta^k}} \frac{C}{V(x, y)} \left( \frac{d(x, y)}{\delta^k} \right)^\omega \left( \frac{d(x, x')}{\delta^k} \right)^\eta \exp(-\nu(\delta^{-k}d(x, y))^a) \\ & \quad + \sum_{\substack{k: \delta^k < (2A_0)^{-1}d(x, y), \\ d(x, x') > \delta^k, d(y, y') > \delta^k}} \frac{C}{V(x, y)} \left( \frac{d(x, y)}{\delta^k} \right)^\omega \exp(-\nu(\delta^{-k}d(x, y))^a) \\ & =: B_1 + B_2 + B_3 + B_4 + B_5. \end{aligned}$$

Following Lemma 8.3 in [AH], an application of the estimate in Remark 2.3 shows that  $B_1$  is bounded by  $\frac{C}{V(x, y)} \left( \frac{d(x, x')}{d(x, y)} \right)^\eta \left( \frac{d(y, y')}{d(x, y)} \right)^\eta$ . Further,  $B_2$  satisfies the same estimate since  $\sum_{m=0}^\infty \delta^{-m(\omega+\eta)} \exp\{-\nu\delta^{-ma}\} \leq C$ . We can deal with  $B_3$ ,  $B_4$  and  $B_5$  similarly. Thus (3.21) is proved.

We remark that the estimate (3.21) is crucial for the proof of (3.12); see the estimate for  $T(f_2)$  below.

Now that we have established the estimates (3.18)–(3.21) on the kernel  $K(x, y)$  of  $T$ , we return to estimating the expressions  $p(x)$  and  $q(x)$ . The size condition on the kernel  $K(x, y)$  and the smoothness condition (3.14) on  $f_1$  yield

$$\begin{aligned} |p(x)| &\leq C \|f\|_{G(\beta, \gamma)} \int_{d(x, y) \leq 4A_0^2 r} \frac{1}{V(x, y)} \left( \frac{d(x, y)}{1 + d(x, x_0)} \right)^\beta \\ &\quad \times \frac{1}{V_1(x_0) + V(x, x_0)} \left( \frac{1}{1 + d(x, x_0)} \right)^\gamma d\mu(y) \\ &\leq C \|f\|_{G(\beta, \gamma)} \left( \frac{d(x, x')}{1 + d(x, x_0)} \right)^\beta \frac{1}{V_1(x_0) + V(x, x_0)} \left( \frac{1}{1 + d(x, x_0)} \right)^\gamma. \end{aligned}$$

This estimate still holds when  $x$  is replaced by  $x'$ , for  $d(x, x') = r$ . Thus

$$|p(x) - p(x')| \leq C \|f\|_{G(\beta, \gamma)} \left( \frac{d(x, x')}{1 + d(x, x_0)} \right)^\beta \frac{1}{V_1(x_0) + V(x, x_0)} \left( \frac{1}{1 + d(x, x_0)} \right)^\gamma.$$

For  $q(x)$ , since  $T1 = 0$  (by the definition of  $D_k(x, y)$  and the cancellation property of  $\psi_\alpha^k$ ), we obtain

$$\begin{aligned} q(x) - q(x') &= \int_X [K(x, y) - K(x', y)] v(y) [f_1(y) - f_1(x)] d\mu(y) \\ &\quad + [f_1(y) - f_1(x)] \int_X K(x, y) u(y) d\mu(y) \\ &=: (E) + (F). \end{aligned}$$

We claim that there exists a constant  $C$  such that for all  $x$ ,

$$(3.23) \quad \left| \int_X K(x, y) u(y) d\mu(y) \right| \leq C.$$

Assuming this claim (which is proved below), together with the estimate for  $f_1$  in (3.14), we find that

$$|(F)| \leq C |f_1(x) - f_1(x')| \leq C \|f\|_{G(\beta, \gamma)} \left( \frac{d(x, x')}{1 + d(x, x_0)} \right)^\beta \frac{1}{V_1(x_0) + V(x, x_0)} \left( \frac{1}{1 + d(x, x_0)} \right)^\gamma.$$

Applying the smoothness estimates for both  $f_1$  and  $K(x, y)$ , we obtain

$$\begin{aligned} |(E)| &\leq C \int_{d(x, y) \geq 4A_0^2 r} |K(x, y) - K(x', y)| |v(y)| |f_1(y) - f_1(x)| d\mu(y) \\ &\leq C \|f\|_{G(\beta, \gamma)} \int_{d(x, y) \geq 4A_0^2 r} \frac{1}{V(x, y)} \left( \frac{d(x, x')}{d(x, y)} \right)^\eta \\ &\quad \times \left( \frac{d(x, y)}{1 + d(x, x_0)} \right)^\beta \frac{1}{V_1(x_0) + V(x, x_0)} \left( \frac{1}{1 + d(x, x_0)} \right)^\gamma d\mu(y) \\ &\leq C \|f\|_{G(\beta, \gamma)} \left( \frac{d(x, x')}{1 + d(x, x_0)} \right)^\beta \frac{1}{V_1(x_0) + V(x, x_0)} \left( \frac{1}{1 + d(x, x_0)} \right)^\gamma, \end{aligned}$$

since  $\beta < \eta$ . Therefore

$$|T(f_1)(x) - T(f_1)(x')| \leq C \|f\|_{G(\beta, \gamma)} \left( \frac{d(x, x')}{1 + d(x, x_0)} \right)^\beta \frac{1}{V_1(x_0) + V(x, x_0)} \left( \frac{1}{1 + d(x, x_0)} \right)^\gamma.$$

We consider three cases. First suppose  $d(x, x') = r \leq (20A_0^2)^{-1}(1 + R)$  and  $R \geq 10$ . Then the points  $x$  and  $x'$  are not in the supports of  $f_2$  and  $f_3$ . Using the double smoothness and

smoothness conditions ((3.21) and (3.19) respectively) on  $K(x, y)$ , and the estimate (3.17) of  $f_2$ , we find

$$\begin{aligned}
|T(f_2)(x) - T(f_2)(x')| &= \left| \int_X [K(x, y) - K(x', y)] f_2(y) d\mu(y) \right| \\
&\leq \int_X |K(x, y) - K(x', y) - K(x, x_0) + K(x', x_0)| |f_2(y)| d\mu(y) \\
&\quad + |K(x, x_0) - K(x', x_0)| \left| \int_X f_2(y) d\mu(y) \right| \\
&\leq C \|f\|_{G(\beta, \gamma)} \left( \frac{d(x, x')}{1 + d(x, x_0)} \right)^\beta \frac{1}{V_1(x_0) + V(x, x_0)} \left( \frac{1}{1 + d(x, x_0)} \right)^\gamma.
\end{aligned}$$

Also,

$$\begin{aligned}
|T(f_3)(x) - T(f_3)(x')| &= \left| \int_X [K(x, y) - K(x', y)] f_3(y) d\mu(y) \right| \\
&\leq C \int_{d(x, y) \geq \frac{R}{8} \geq 2A_0 r} \frac{1}{V(x, y)} \left( \frac{d(x, x')}{d(x, y)} \right)^\eta |f_3(y)| d\mu(y) \\
&\leq C \|f\|_{G(\beta, \gamma)} \left( \frac{d(x, x')}{1 + d(x, x_0)} \right)^\beta \frac{1}{V_1(x_0) + V(x, x_0)} \left( \frac{1}{1 + d(x, x_0)} \right)^\gamma.
\end{aligned}$$

In the second case, where  $d(x, x_0) = R$  and  $(2A_0)^{-1}(1 + R) \geq d(x, x') = r \geq (20A_0^2)^{-1}(1 + R)$ , the desired estimate for  $T(f)(x)$  follows from the estimate of (3.6). So we need only consider the third case, where  $R \leq 10$  and  $r \leq 11/(20A_0)$ . This case is similar, and indeed easier. In fact, all we need to do is to replace  $R$  in the proof above by 10. We leave the details to the reader. This completes the proof of (3.7).

To finish the argument for Theorem 3.4, it remains to establish the claim (3.23). To do so, we prove that there exists a constant  $C$  such that

$$(3.24) \quad \|T\phi\|_\infty \leq C$$

for all functions  $\phi$  with the properties that  $\|\phi\|_\infty \leq 1$  and there exist  $x_0 \in X$  and  $t > 0$  such that  $\text{supp } \phi \subseteq B(x_0, t)$  and  $\|\phi\|_\eta := \sup_{x \neq y} \{|\phi(x) - \phi(y)|/d(x, y)^\eta\} \leq t^{-\eta}$ .

We again follow the idea of Meyer's proof in [M2]. Let  $\chi_0(x) = h(x)$ , where  $h(x)$  is a smooth cut-off function as in Lemma 3.8 with the property that  $h(x) \equiv 1$  on  $B(x_0, 2t)$  and  $h(x) \equiv 0$  on  $B(x_0, 8A_0^2 t)^c$ . Set  $\chi_1 := 1 - \chi_0$ . Then  $\phi = \phi\chi_0$  and for all  $\psi \in C_0^\eta(X)$ ,

$$\begin{aligned}
\langle T\phi, \psi \rangle &= \langle K(x, y), \phi(y)\psi(x) \rangle = \langle K(x, y), \chi_0(y)\phi(y)\psi(x) \rangle \\
&= \langle K(x, y), \chi_0(y)[\phi(y) - \phi(x)]\psi(x) \rangle + \langle K(x, y), \chi_0(y)\phi(x)\psi(x) \rangle \\
&:= (G) + (H).
\end{aligned}$$

Applying the size condition (3.18) on the kernel  $K(x, y)$  yields

$$|(G)| \leq C\|\psi\|_1.$$

To estimate (H), it suffices to show that for  $x \in B(x_0, t)$ ,

$$(3.25) \quad |T\chi_0(x)| \leq C,$$

since as  $(H) = \langle T\chi_0, \phi\psi \rangle$ , we then have

$$|(H)| \leq \|T\chi_0\|_{L^\infty(B(x_0, t))} \|\phi\psi\|_{L^1(B(x_0, t))} \leq C\|\psi\|_1.$$

To show (3.25), we use Meyer's idea again [M2]. Take  $\psi \in C^\eta(X)$  with  $\text{supp } \psi \subseteq B(x_0, t)$  and  $\int_X \psi(x) d\mu(x) = 0$ . Since  $T1 = 0$  and  $\int_X \psi(x) d\mu(x) = 0$ , and using the smoothness condition (3.19) on  $K(x, y)$ , we obtain

$$\begin{aligned} |\langle T\chi_0, \psi \rangle| &= |-\langle T\chi_1, \psi \rangle| = \left| \iint_{X \times X} [K(x, y) - K(x_0, y)] \chi_1(y) \psi(x) d\mu(x) d\mu(y) \right| \\ &\leq C \|\psi\|_1. \end{aligned}$$

Thus,  $T\chi_0(x) = \Lambda + \gamma(x)$  for  $x \in B(x_0, t)$ , where  $\Lambda$  is a constant and  $\|\gamma\|_\infty \leq C$ . To estimate  $\Lambda$ , choose  $\phi_1 \in C_0^\eta(X)$  with  $\text{supp } \phi_1 \subseteq B(x_0, t)$ ,  $\|\phi_1\|_\infty \leq 1$ ,  $\|\phi_1\|_\eta \leq t^{-\eta}$  and  $\int_X \phi_1(x) d\mu(x) = Ct$ . Since  $T$  is bounded on  $L^2(X)$ , we have

$$\left| Ct\Lambda + \int_X \phi_1(x) \gamma(x) d\mu(x) \right| = |\langle T\chi_0, \phi_1 \rangle| \leq Ct.$$

Therefore  $|\Lambda| \leq C$ , and hence the claim (3.23) is proved. This completes the proof of Theorem 3.4, modulo the proof of Lemma 3.6.  $\square$

It remains to prove the technical lemma used in the proof of Theorem 3.4.

*Proof of Lemma 3.6.* (i) To establish the decay condition (3.9), we write

$$|D_k(x, y)| = \left| \sum_{\alpha \in \mathcal{Y}^k} \mu(B(y_\alpha^k, \delta^k)) \frac{\psi_\alpha^k(x)}{\sqrt{\mu(B(y_\alpha^k, \delta^k))}} \frac{\psi_\alpha^k(y)}{\sqrt{\mu(B(y_\alpha^k, \delta^k))}} \right|.$$

By Theorem 3.3, we know that  $\psi_\alpha^k(x)/\sqrt{\mu(B(y_\alpha^k, \delta^k))}$  belongs to  $G(y_\alpha^k, \delta^k, \eta, \gamma + \eta)$ . Applying the size condition (i) from Definition 3.1, we see that

$$\begin{aligned} |D_k(x, y)| &\leq C \sum_{\alpha \in \mathcal{Y}^k} \mu(B(y_\alpha^k, \delta^k)) \frac{1}{V_{\delta^k}(y_\alpha^k) + V(y_\alpha^k, x)} \left( \frac{\delta^k}{\delta^k + d(y_\alpha^k, x)} \right)^{\gamma + \eta} \\ &\quad \times \frac{1}{V_{\delta^k}(y_\alpha^k) + V(y_\alpha^k, y)} \left( \frac{\delta^k}{\delta^k + d(y_\alpha^k, y)} \right)^{\gamma + \eta}. \end{aligned}$$

Note that for each  $z \in B(y_\alpha^k, \delta^k)$  one can replace  $y_\alpha^k$  by  $z$  to get  $\delta^k + d(y_\alpha^k, x) \sim \delta^k + d(z, x)$  and  $V_{\delta^k}(y_\alpha^k) + V(y_\alpha^k, x) \sim V_{\delta^k}(z) + V(z, x)$ , and similarly for  $\delta^k + d(y_\alpha^k, y)$  and  $V_{\delta^k}(y_\alpha^k) + V(y_\alpha^k, y)$ . Thus, first replacing  $\mu(B(y_\alpha^k, \delta^k))$  by  $\int_{B(y_\alpha^k, \delta^k)} d\mu(z)$  and then replacing  $y_\alpha^k$  by  $z$ , and finally summing up over  $\alpha \in \mathcal{Y}^k$ , we find that the last sum above is bounded by

$$\begin{aligned} &C \int_X \frac{1}{V_{\delta^k}(z) + V(z, x)} \left( \frac{\delta^k}{\delta^k + d(z, x)} \right)^{\gamma + \eta} \frac{1}{V_{\delta^k}(z) + V(z, y)} \left( \frac{\delta^k}{\delta^k + d(z, y)} \right)^{\gamma + \eta} d\mu(z) \\ &=: (P) + (Q). \end{aligned}$$

Here (P) is the result of integrating over the set  $d(x, z) \leq (2A_0)^{-1}(\delta^k + d(x, y))$  and (Q) over the set  $d(x, z) > (2A_0)^{-1}(\delta^k + d(x, y))$ . To estimate (P), note that if  $d(x, z) \leq (2A_0)^{-1}(\delta^k + d(x, y))$  and  $2\delta^k \leq d(x, y)$ , then  $d(y, z) > (10A_0)^{-1}(\delta^k + d(x, y))$  and by the doubling property,

$$V(z, y) = \mu(B(y, d(z, y))) \geq \mu(B(y, (10A_0)^{-1}d(x, y))) \geq (10A_0)^{-\omega} V(x, y).$$

Therefore,  $V_{\delta^k}(x) + V(x, y) \leq C(V_{\delta^k}(z) + V(z, y))$ . Next, if  $d(x, z) \leq (2A_0)^{-1}(\delta^k + d(x, y))$  and  $2\delta^k > d(x, y)$ , then  $d(x, z) \leq 3(2A_0)^{-1}\delta^k$ . For this case, first suppose that  $d(x, z) \leq \delta^k$  and hence,  $V_{\delta^k}(z) \sim V_{\delta^k}(x)$ . On the other hand, if  $d(x, z) > \delta^k$ , then  $(2A_0 - 1)\delta^k \leq (2A_0)^{-1}d(x, y)$

and hence  $V_{\delta^k}(x) \leq CV(x, y)$ . Therefore, in this case, again  $V_{\delta^k}(x) + V(x, y) \leq C(V_{\delta^k}(z) + V(z, y))$  and thus we get

$$(P) \leq C \frac{1}{V_{\delta^k}(x) + V(x, y)} \left( \frac{\delta^k}{\delta^k + d(x, y)} \right)^{\gamma+\eta} \leq C \frac{1}{V_{\delta^k}(x) + V(x, y)} \left( \frac{\delta^k}{\delta^k + d(x, y)} \right)^{\gamma},$$

as required. The estimate for (Q) is the same, but with  $x$  and  $y$  reversed.

(ii) To establish the smoothness condition (3.10), we write

$$\begin{aligned} & |D_k(x, y) - D(x, y')| \\ &= \sum_{\alpha \in \mathcal{Y}^k} \mu(B(y_\alpha^k, \delta^k)) \left| \frac{\psi_\alpha^k(x)}{\sqrt{\mu(B(y_\alpha^k, \delta^k))}} \left[ \frac{\psi_\alpha^k(y)}{\sqrt{\mu(B(y_\alpha^k, \delta^k))}} - \frac{\psi_\alpha^k(y')}{\sqrt{\mu(B(y_\alpha^k, \delta^k))}} \right] \right| \\ &=: (R) + (S). \end{aligned}$$

Here (R) is the result of summing over the set of  $\alpha \in \mathcal{Y}^k$  such that  $d(y, y') \leq (2A_0)^{-1}(\delta^k + d(y, y_\alpha^k))$  or  $d(y, y') \leq (2A_0)^{-1}(\delta^k + d(y', y_\alpha^k))$ , and (S) over the set of  $\alpha \in \mathcal{Y}^k$  such that  $d(y, y') > (2A_0)^{-1}(\delta^k + d(y, y_\alpha^k))$  and  $d(y, y') > (2A_0)^{-1}(\delta^k + d(y', y_\alpha^k))$ .

For (R), use the size condition (Definition 3.1(i)) for the first factor  $\psi_\alpha^k(x)/\sqrt{\mu(B(y_\alpha^k, \delta^k))}$  and the Hölder regularity condition (Definition 3.1(ii)) for the terms  $\psi_\alpha^k(y)/\sqrt{\mu(B(y_\alpha^k, \delta^k))}$  in the second factor. We find that

$$\begin{aligned} (R) &\leq C \sum_{\alpha \in \mathcal{Y}^k} \mu(B(y_\alpha^k, \delta^k)) \left( \frac{d(y, y')}{\delta^k + d(y_\alpha^k, y)} \right)^\eta \frac{1}{V_{\delta^k}(y_\alpha^k) + V(y_\alpha^k, y)} \left( \frac{\delta^k}{\delta^k + d(y_\alpha^k, y)} \right)^{\gamma+\eta} \\ &\quad \times \frac{1}{V_{\delta^k}(y_\alpha^k) + V(y_\alpha^k, x)} \left( \frac{\delta^k}{\delta^k + d(y_\alpha^k, x)} \right)^{\gamma+\eta}. \end{aligned}$$

Applying the same proof as for (3.9), we see that the last sum above is bounded by

$$\begin{aligned} & C \int_X \left( \frac{d(y, y')}{\delta^k + d(z, y)} \right)^\eta \frac{1}{V_{\delta^k}(z) + V(z, y)} \left( \frac{\delta^k}{\delta^k + d(z, y)} \right)^{\gamma+\eta} \\ & \quad \times \frac{1}{V_{\delta^k}(z) + V(z, x)} \left( \frac{\delta^k}{\delta^k + d(z, x)} \right)^{\gamma+\eta} d\mu(z) \\ & \leq C \left( \frac{d(y, y')}{\delta^k + d(x, y)} \right)^\eta \frac{1}{V_{\delta^k}(x) + V(x, y)} \left( \frac{\delta^k}{\delta^k + d(x, y)} \right)^\gamma. \end{aligned}$$

To deal with (S), we can write

$$\begin{aligned} (S) &\leq \sum_{\alpha \in \mathcal{Y}^k: d(y, y') > (2A_0)^{-1}(\delta^k + d(y, y_\alpha^k))} \mu(B(y_\alpha^k, \delta^k)) \left| \frac{\psi_\alpha^k(x)}{\sqrt{\mu(B(y_\alpha^k, \delta^k))}} \frac{\psi_\alpha^k(y)}{\sqrt{\mu(B(y_\alpha^k, \delta^k))}} \right| \\ &\quad + \sum_{\alpha \in \mathcal{Y}^k: d(y, y') > (2A_0)^{-1}(\delta^k + d(y', y_\alpha^k))} \mu(B(y_\alpha^k, \delta^k)) \left| \frac{\psi_\alpha^k(x)}{\sqrt{\mu(B(y_\alpha^k, \delta^k))}} \frac{\psi_\alpha^k(y')}{\sqrt{\mu(B(y_\alpha^k, \delta^k))}} \right|. \end{aligned}$$

For the first sum, following the same approach as for (R) but with  $d(y, y') > (2A_0)^{-1}(\delta^k + d(y, y_\alpha^k))$ , we must deal with the integral

$$\int_X \left( \frac{d(y, y')}{\delta^k + d(z, y)} \right)^\eta \frac{1}{V_{\delta^k}(z) + V(z, y)} \left( \frac{\delta^k}{\delta^k + d(z, y)} \right)^{\gamma+\eta} \frac{1}{V_{\delta^k}(z) + V(z, x)} \left( \frac{\delta^k}{\delta^k + d(z, x)} \right)^{\gamma+\eta} d\mu(z).$$



Applying the same proof as for (R), but using the size condition (Definition 3.1(i)) for both factors, we obtain that this integral is bounded by

$$C \left( \frac{d(y, y')}{\delta^k + d(x, y)} \right)^\eta \frac{1}{V_{\delta^k}(x) + V(x, y)} \left( \frac{\delta^k}{\delta^k + d(x, y)} \right)^\gamma.$$

The second sum is similar to the first one, with  $y$  and  $y'$  reversed. Thus, by the same proof we find that the second sum is bounded by

$$C \left( \frac{d(y, y')}{\delta^k + d(x, y')} \right)^\eta \frac{1}{V_{\delta^k}(x) + V(x, y')} \left( \frac{\delta^k}{\delta^k + d(x, y')} \right)^\gamma.$$

Note that the fact  $d(y, y') \leq (2A_0)^{-1}(\delta^k + d(x, y))$  implies that  $\delta^k + d(x, y) \sim \delta^k + d(x, y')$  and  $V_{\delta^k}(x) + V(x, y) \sim V_{\delta^k}(x) + V(x, y')$ . Therefore, we obtain the desired estimate for the second sum.

(iii) The proof for the double smoothness condition (3.11) is similar to that for (3.10), and we omit the details.

This completes the proof of Lemma 3.6.  $\square$

**3.2. Product test functions, distributions, and wavelet reproducing formula.** We now consider the product setting  $(X_1, d_1, \mu_1) \times (X_2, d_2, \mu_2)$ , where  $(X_i, d_i, \mu_i)$ ,  $i = 1, 2$ , are spaces of homogeneous type as defined in the Introduction. For  $i = 1, 2$ , let  $C_{\mu_i}$  be the doubling constant as in inequality (1.2), let  $\omega_i$  be the upper dimension as in inequality (1.3), and let  $A_0^{(i)}$  be the constant in the quasi-triangle inequality (1.1). In this subsection we use the notation  $(x, y)$  for an element of  $X_1 \times X_2$ .

On each  $X_i$  there is a wavelet basis  $\{\psi_{\alpha_i}^{k_i}\}$ , with Hölder exponent  $\eta_i$  as in inequality (2.11).

We now define the spaces of test functions and distributions on the product space  $X_1 \times X_2$ .

**Definition 3.9.** (Product test functions) Let  $(x_0, y_0) \in X_1 \times X_2$  and  $r = (r_1, r_2)$  with  $r_1, r_2 > 0$ . Take  $\beta = (\beta_1, \beta_2)$ , with  $\beta_1 \in (0, \eta_1]$ ,  $\beta_2 \in (0, \eta_2]$ , and  $\gamma = (\gamma_1, \gamma_2)$  with  $\gamma_1, \gamma_2 > 0$ . A function  $f(x, y)$  defined on  $X_1 \times X_2$  is said to be a *test function of type*  $(x_0, y_0; r; \beta; \gamma)$  if the following three conditions hold.

- (a) For each fixed  $y \in X_2$ ,  $f(x, y)$  as a function of the variable  $x \in X_1$  is a test function in  $G(x_0, r_1, \beta_1, \gamma_1)$ .
- (b) For each fixed  $x \in X_1$ ,  $f(x, y)$  as a function of the variable  $y \in X_2$  is a test function in  $G(y_0, r_2, \beta_2, \gamma_2)$ .
- (c) The following properties hold:
  - (i) (Size condition) For all  $y \in X_2$ ,

$$\|f(\cdot, y)\|_{G(x_0, r_1, \beta_1, \gamma_1)} \leq C \frac{1}{V_{r_2}(y_0) + V(y, y_0)} \left( \frac{r_2}{r_2 + d_2(y, y_0)} \right)^{\gamma_2}.$$

- (ii) (Hölder regularity condition) For all  $y, y' \in X_2$  such that  $d_2(y, y') \leq (2A_0^{(2)})^{-1}(r_2 + d_2(y, y_0))$ , we have

$$\|f(\cdot, y) - f(\cdot, y')\|_{G(x_0, r_1, \beta_1, \gamma_1)} \leq C \left( \frac{d_2(y, y')}{r_2 + d_2(y, y_0)} \right)^{\beta_2} \frac{1}{V_{r_2}(y_0) + V(y_0, y)} \left( \frac{r_2}{r_2 + d_2(y, y_0)} \right)^{\gamma_2}.$$

- (iii) Properties (i) and (ii) also hold with  $x$  and  $y$  interchanged.

- (iv) (Cancellation condition)  $\int_{X_1} f(x, y) d\mu_1(x) = 0$  for all  $y \in X_2$ , and  $\int_{X_2} f(x, y) d\mu_2(y) = 0$  for all  $x \in X_1$ .

When  $f$  is a test function of type  $(x_0, y_0; r; \beta; \gamma)$ , we write  $f \in G(x_0, y_0; r; \beta; \gamma)$ . Note the use of semicolons here to distinguish the product definition from the one-parameter version.

The expression

$$\|f\|_{G(x_0, y_0; r; \beta; \gamma)} := \inf\{C : \text{(i), (ii) and (iii) hold}\}$$

defines a norm on  $G(x_0, y_0; r; \beta; \gamma)$ .

We denote by  $G(\beta_1, \beta_2; \gamma_1, \gamma_2)$  the class  $G(x_0, y_0; 1, 1; \beta; \gamma)$  for arbitrary fixed  $(x_0, y_0) \in X_1 \times X_2$ . Then  $G(x_0, y_0; r; \beta; \gamma) = G(\beta_1, \beta_2; \gamma_1, \gamma_2)$ , with equivalent norms, for all  $(x_0, y_0) \in X_1 \times X_2$  and  $r_1, r_2 > 0$ . Furthermore,  $G(\beta_1, \beta_2; \gamma_1, \gamma_2)$  is a Banach space with respect to the norm on  $G(\beta_1, \beta_2; \gamma_1, \gamma_2)$ .

For  $\beta_i \in (0, \eta_i]$  and  $\gamma_i > 0$ , for  $i = 1, 2$ , let  $\mathring{G}(\beta_1, \beta_2; \gamma_1, \gamma_2)$  be the completion of the space  $G(\eta_1, \eta_2; \gamma_1, \gamma_2)$  in  $G(\beta_1, \beta_2; \gamma_1, \gamma_2)$  in the norm of  $G(\beta_1, \beta_2; \gamma_1, \gamma_2)$ . We define the norm on  $\mathring{G}(\beta_1, \beta_2; \gamma_1, \gamma_2)$  by  $\|f\|_{\mathring{G}(\beta_1, \beta_2; \gamma_1, \gamma_2)} := \|f\|_{G(\beta_1, \beta_2; \gamma_1, \gamma_2)}$ .

The (scaled) product wavelets given by  $\psi_{\alpha_1}^{k_1}(x)\psi_{\alpha_1}^{k_1}(y)(\mu_1(B(y_{\alpha_1}^{k_1}, \delta_1^{k_1}))\mu_2(B(y_{\alpha_2}^{k_2}, \delta_2^{k_2})))^{-1/2}$  are product test functions in  $G(y_{\alpha_1}^{k_1}, y_{\alpha_2}^{k_2}; \delta; \beta; \gamma)$  for each  $\gamma = (\gamma_1, \gamma_2)$  with  $\gamma_1 > 0, \gamma_2 > 0$ , where  $\delta = (\delta_1^{k_1}, \delta_2^{k_2})$  and  $\beta = (\eta_1, \eta_2)$ ; this is straightforward to check.

**Definition 3.10.** (Product distributions) Let  $(x_0, y_0) \in X_1 \times X_2$  and  $r = (r_1, r_2)$  with  $r_1, r_2 > 0$ . Take  $\beta = (\beta_1, \beta_2)$ , with  $\beta_1 \in (0, \eta_1], \beta_2 \in (0, \eta_2]$ , and  $\gamma = (\gamma_1, \gamma_2)$  with  $\gamma_1, \gamma_2 > 0$ . We define the *distribution space*  $(\mathring{G}(\beta_1, \beta_2; \gamma_1, \gamma_2))'$  to consist of all linear functionals  $\mathcal{L}$  from  $\mathring{G}(\beta_1, \beta_2; \gamma_1, \gamma_2)$  to  $\mathbb{C}$  with the property that there exists a constant  $C$  such that for all  $f \in \mathring{G}(\beta_1, \beta_2; \gamma_1, \gamma_2)$ ,

$$|\mathcal{L}(f)| \leq C\|f\|_{\mathring{G}(\beta_1, \beta_2; \gamma_1, \gamma_2)}.$$

We have the following version of the wavelet reproducing formula in the product setting  $X_1 \times X_2$ .

**Theorem 3.11.** (Product reproducing formula) Take  $\beta_i, \gamma_i \in (0, \eta_i)$  for  $i = 1, 2$ .

(a) *The wavelet reproducing formula*

$$(3.26) \quad f(x, y) = \sum_{k_1} \sum_{\alpha_1 \in \mathcal{D}^{k_1}} \sum_{k_2} \sum_{\alpha_2 \in \mathcal{D}^{k_2}} \langle f, \psi_{\alpha_1}^{k_1} \psi_{\alpha_2}^{k_2} \rangle \psi_{\alpha_1}^{k_1}(x) \psi_{\alpha_2}^{k_2}(y)$$

holds in the space of test functions  $\mathring{G}(\beta'_1, \beta'_2; \gamma'_1, \gamma'_2)$  for each  $\beta'_i \in (0, \beta_i)$  and  $\gamma'_i \in (0, \gamma_i)$ , for  $i = 1, 2$ .

(b) *The wavelet reproducing formula (3.26) also holds in the space of distributions  $(\mathring{G}(\beta_1, \beta_2; \gamma_1, \gamma_2))'$ .*

*Proof.* As before, the wavelet reproducing formula for distributions follows immediately from that for test functions. The proof for test functions proceeds by iteration of Theorem 3.4. Write

$$\begin{aligned} g(x, y) &:= \sum_{|k_1| \leq L_1} \sum_{\alpha_1 \in \mathcal{D}^{k_1}} \sum_{|k_2| \leq L_2} \sum_{\alpha_2 \in \mathcal{D}^{k_2}} \langle f, \psi_{\alpha_1}^{k_1} \psi_{\alpha_2}^{k_2} \rangle \psi_{\alpha_1}^{k_1}(x) \psi_{\alpha_2}^{k_2}(y) - f(x, y) \\ &=: g_1(x, y) + g_2(x, y), \end{aligned}$$

where

$$\begin{aligned} g_1(x, y) &:= \sum_{|k_1| \leq L_1} \sum_{\alpha_1 \in \mathcal{Y}^{k_1}} \left\langle \psi_{\alpha_1}^{k_1}, \sum_{|k_2| \leq L_2} \sum_{\alpha_2 \in \mathcal{Y}^{k_2}} \langle f(\cdot, \cdot), \psi_{\alpha_2}^{k_2} \rangle \psi_{\alpha_2}^{k_2}(y) \right\rangle \psi_{\alpha_1}^{k_1}(x) \\ &\quad - \sum_{|k_2| \leq L_2} \sum_{\alpha_2 \in \mathcal{Y}^{k_2}} \langle f(x, \cdot), \psi_{\alpha_2}^{k_2} \rangle \psi_{\alpha_2}^{k_2}(y) \end{aligned}$$

and

$$g_2(x, y) := \sum_{|k_2| \leq L_2} \sum_{\alpha_2 \in \mathcal{Y}^{k_2}} \langle f(x, \cdot), \psi_{\alpha_2}^{k_2} \rangle \psi_{\alpha_2}^{k_2}(y) - f(x, y).$$

To see the convergence in the space of test functions, we recall the following (one-parameter) estimate on  $X$ , as shown in the proof of Theorem 3.4: Given  $\beta, \gamma \in (0, \eta)$ , for each  $\beta' \in (0, \beta)$  and  $\gamma' \in (0, \gamma)$  there is a constant  $\sigma > 0$  such that for each positive integer  $L$

$$(3.27) \quad \left\| f(\cdot) - \sum_{|k| \leq L} \sum_{\alpha \in \mathcal{Y}^k} \langle \psi_{\alpha}^k, f \rangle \psi_{\alpha}^k(\cdot) \right\|_{G(\beta', \gamma')} \leq C \delta^{\sigma L} \|f\|_{G(\beta, \gamma)},$$

where  $C$  is a constant independent of  $f \in \mathring{G}(\beta, \gamma)$ . Note that inequality (3.27) is the same as inequality (3.5), slightly rewritten. Inequality (3.27), together with the triangle inequality, implies that

$$(3.28) \quad \left\| \sum_{|k| \leq L} \sum_{\alpha \in \mathcal{Y}^k} \langle \psi_{\alpha}^k, f \rangle \psi_{\alpha}^k(\cdot) \right\|_{G(\beta', \gamma')} \leq C \|f\|_{G(\beta, \gamma)}.$$

We observe that if  $f \in G(\beta_1, \beta_2; \gamma_1, \gamma_2)$ , then  $\|f(\cdot, y)\|_{G(\beta_1, \gamma_1)}$ , as a function of the variable  $y$ , is in  $G(\beta_2, \gamma_2)$ , and satisfies  $\| \|f(\cdot, \cdot)\|_{G(\beta_1, \gamma_1)} \|_{G(\beta_2, \gamma_2)} \leq \|f\|_{G(\beta_1, \beta_2; \gamma_1, \gamma_2)}$ . Similarly,  $\| \|f(\cdot, \cdot)\|_{G(\beta_2, \gamma_2)} \|_{G(\beta_1, \gamma_1)} \leq \|f\|_{G(\beta_1, \beta_2; \gamma_1, \gamma_2)}$ . Therefore, we obtain

$$\begin{aligned} \|g_1(\cdot, y)\|_{G(\beta'_1, \gamma'_1)} &\leq C \delta^{L_1 \sigma} \left\| \sum_{|k_2| \leq L_2} \sum_{\alpha_2 \in \mathcal{Y}^{k_2}} \langle f(\cdot, \cdot), \psi_{\alpha_2}^{k_2} \rangle \psi_{\alpha_2}^{k_2}(y) \right\|_{G(\beta_1, \gamma_1)} \\ &\leq C \delta^{L_1 \sigma} \|f(\cdot, \cdot)\|_{G(\beta_2, \gamma_2)} \frac{1}{V_{r_2}(y_0) + V(y_0, y)} \left( \frac{r_2}{r_2 + d(y, y_0)} \right)^{\gamma_2} \|_{G(\beta_1, \gamma_1)} \\ &\leq C \delta^{L_1 \sigma} \|f(\cdot, \cdot)\|_{G(\beta_1, \beta_2; \gamma_1, \gamma_2)} \frac{1}{V_{r_2}(y_0) + V(y_0, y)} \left( \frac{r_2}{r_2 + d(y, y_0)} \right)^{\gamma_2}, \end{aligned}$$

where the first inequality follows from (3.27) and the second inequality follows from (3.28). Similarly,

$$\|g_2(x, y)\|_{G(\beta'_1, \gamma'_1)} \leq C \delta^{L_2 \sigma} \|f\|_{G(\beta_1, \beta_2; \gamma_1, \gamma_2)} \frac{1}{V_{r_2}(y_0) + V(y_0, y)} \left( \frac{r_2}{r_2 + d(y, y_0)} \right)^{\gamma_2}.$$

Noting that  $g(x, y) - g(x, y') = [g_1(x, y) - g_1(x, y')] + [g_2(x, y) - g_2(x, y')]$ , by repeating the same estimates we obtain

$$\begin{aligned} \|g(\cdot, y) - g(\cdot, y')\|_{G(\beta'_1, \gamma'_1)} &\leq C(\delta^{L_1 \sigma} + \delta^{L_2 \sigma}) \|f(\cdot, \cdot)\|_{G(\beta_1, \beta_2; \gamma_1, \gamma_2)} \\ &\quad \times \left( \frac{d(y, y')}{r_2 + d(y, y_0)} \right)^{\beta_2} \frac{1}{V_{r_2}(y_0) + V(y_0, y)} \left( \frac{r_2}{r_2 + d(y, y_0)} \right)^{\gamma_2} \end{aligned}$$

where  $d(y, y') \leq (2A_0^{(2)})^{-1}(r_2 + d(y, y_0))$ .

The same proof can be carried out for the estimates with  $x$  and  $y$  interchanged. Hence

$$\|g\|_{G(\beta'_1, \beta'_2; \gamma'_1, \gamma'_2)} \leq C(\delta^{L_1\sigma} + \delta^{L_2\sigma})\|f\|_{G(\beta_1, \beta_2; \gamma_1, \gamma_2)},$$

which yields the convergence in  $\mathring{G}(\beta'_1, \beta'_2; \gamma'_1, \gamma'_2)$ .  $\square$

#### 4. LITTLEWOOD–PALEY SQUARE FUNCTIONS AND PLANCHEREL–PÓLYA INEQUALITIES

We now carry out the philosophy described near the end of the introduction, in order to establish the Littlewood–Paley theory for the discrete square function in terms of wavelet coefficients. We define the discrete and continuous square functions, and prove their norm-equivalence via Plancherel–Pólya inequalities, whose proof takes up most of this section. Again we begin with the one-parameter case.

**4.1. One-parameter square functions via wavelets, and Plancherel–Pólya inequalities.** We first apply the orthonormal wavelet basis constructed in [AH] to introduce the discrete Littlewood–Paley square function, defined via the wavelet coefficients as follows.

**Definition 4.1.** (Discrete square function in terms of wavelet coefficients) For  $f$  in  $(\mathring{G}(\beta, \gamma))'$  with  $\beta, \gamma \in (0, \eta)$ , the *discrete Littlewood–Paley square function*  $S(f)$  of  $f$  is defined by

$$(4.1) \quad S(f)(x) := \left\{ \sum_k \sum_{\alpha \in \mathcal{D}^k} |\langle \psi_\alpha^k, f \rangle \tilde{\chi}_{Q_\alpha^k}(x)|^2 \right\}^{1/2},$$

where  $\tilde{\chi}_{Q_\alpha^k}(x) := \chi_{Q_\alpha^k}(x) \mu(Q_\alpha^k)^{-1/2}$  and  $\chi_{Q_\alpha^k}(x)$  is the indicator function of the dyadic cube  $Q_\alpha^k$ .

It is straightforward that  $\|S(f)\|_{L^2(X)} = \|f\|_{L^2(X)}$ , since  $\{\psi_\alpha^k\}$  forms an orthonormal wavelet basis for  $L^2(X)$ . However, it is not easy to see why  $\|S(f)\|_{L^p(X)} \sim \|f\|_{L^p(X)}$  for  $1 < p < \infty$  with  $p \neq 2$ . This difficulty is because the classical method, namely the vector-valued Calderón–Zygmund operator theory, cannot be carried out here due to the lack of smoothness in the  $x$  variable. For this reason, we introduce the following continuous Littlewood–Paley square function in terms of the wavelet operators  $D_k$ .

**Definition 4.2.** (Continuous square function in terms of wavelet operators) Let  $D_k$  be the operator with kernel  $D_k(x, y) = \sum_{\alpha \in \mathcal{D}^k} \psi_\alpha^k(x) \psi_\alpha^k(y)$ . For  $f \in (\mathring{G}(\beta, \gamma))'$  with  $\beta, \gamma \in (0, \eta)$ , the *continuous Littlewood–Paley square function*  $S_c(f)$  of  $f$  is defined by

$$S_c(f)(x) := \left\{ \sum_k |D_k(f)(x)|^2 \right\}^{1/2}.$$

The two main results in this subsection are as follows.

**Theorem 4.3.** (Littlewood–Paley theory) Suppose  $\beta, \gamma \in (0, \eta)$  and  $\frac{\omega}{\omega+\eta} < p < \infty$ , where  $\omega$  is the upper dimension of  $(X, d, \mu)$ . For  $f$  in  $(\mathring{G}(\beta, \gamma))'$ , we have

$$\|S(f)\|_{L^p(X)} \sim \|S_c(f)\|_{L^p(X)}.$$

Moreover, if  $1 < p < \infty$ , then

$$\|S(f)\|_{L^p(X)} \sim \|S_c(f)\|_{L^p(X)} \sim \|f\|_{L^p(X)}.$$

The key idea in proving Theorem 4.3 is the following Plancherel–Pólya type inequalities.

**Theorem 4.4.** (Plancherel–Pólya inequalities) *Suppose  $\beta, \gamma \in (0, \eta)$  and  $\frac{\omega}{\omega+\eta} < p < \infty$ , where  $\omega$  is the upper dimension of  $(X, d, \mu)$ . Fix  $N \in \mathbb{N}$ . Then there is a positive constant  $C$  such that for all  $f \in (\mathring{G}(\beta, \gamma))'$ , we have*

$$(4.2) \quad \left\| \left\{ \sum_{k'} \sum_{\alpha' \in \mathcal{X}^{k'+N}} \left[ \sup_{z \in Q_{\alpha'}^{k'+N}} |D_{k'}(f)(z)|^2 \right] \chi_{Q_{\alpha'}^{k'+N}}(\cdot) \right\}^{1/2} \right\|_{L^p(X)} \leq C \left\| \left\{ \sum_k \sum_{\alpha \in \mathcal{Y}^k} |\langle \psi_\alpha^k, f \rangle \tilde{\chi}_{Q_\alpha^k}(\cdot)|^2 \right\}^{1/2} \right\|_{L^p(X)}.$$

Moreover, for a fixed sufficiently large integer  $N$  ( $N$  will be determined later in the proof), there is a positive constant  $C$  such that for all  $f \in (\mathring{G}(\beta, \gamma))'$ , we have

$$(4.3) \quad \left\| \left\{ \sum_k \sum_{\alpha \in \mathcal{Y}^k} |\langle \psi_\alpha^k, f \rangle \tilde{\chi}_{Q_\alpha^k}(\cdot)|^2 \right\}^{1/2} \right\|_{L^p(X)} \leq C \left\| \left\{ \sum_{k'} \sum_{\alpha' \in \mathcal{X}^{k'+N}} \left[ \inf_{z \in Q_{\alpha'}^{k'+N}} |D_{k'}(f)(z)|^2 \right] \chi_{Q_{\alpha'}^{k'+N}}(\cdot) \right\}^{1/2} \right\|_{L^p(X)}.$$

Note that in each of the inequalities (4.2) and (4.3), on one side, for each  $k \in \mathbb{Z}$  the sum runs over the set  $\mathcal{Y}^k$ , while on the other side for each  $k' \in \mathbb{Z}$  the sum runs over the set  $\mathcal{X}^{k'+N}$ . Besides the distinction between  $\mathcal{Y}$  and  $\mathcal{X}$ , the other difference here is that in the expressions involving  $D_{k'}$ , it is not sufficient to sum at the scale of  $k'$ , but rather, following [DH], we must sum over all cubes at the smaller scale  $k' + N$ .

*Proof of Theorem 4.3.* Theorem 4.3 follows from Theorem 4.4, by standard arguments that can be found in [DH]. We sketch the idea. The first estimate in Theorem 4.3 follows from Theorem 4.4 together with the following observation:

$$\begin{aligned} \sum_{k'} \sum_{\alpha' \in \mathcal{X}^{k'+N}} \inf_{z \in Q_{\alpha'}^{k'+N}} |D_{k'}(f)(z)|^2 \chi_{Q_{\alpha'}^{k'+N}}(x) &\leq \sum_k |D_k(f)(x)|^2 \\ &\leq \sum_{k'} \sum_{\alpha' \in \mathcal{X}^{k'+N}} \sup_{z \in Q_{\alpha'}^{k'+N}} |D_{k'}(f)(z)|^2 \chi_{Q_{\alpha'}^{k'+N}}(x). \end{aligned}$$

For the second estimate in Theorem 4.3, when  $1 < p < \infty$  one obtains from the classical method of vector-valued Calderón–Zygmund operator theory that  $\|S_c(f)\|_{L^p(X)} \leq C\|f\|_{L^p(X)}$ . This estimate together with the wavelet expansion as in (2.13) gives  $\|f\|_{L^p(X)} \leq C\|S_c(f)\|_{L^p(X)}$ , and Theorem 4.3 follows.  $\square$

We would like to point out that to consider  $S_c(f)$  as a vector-valued Calderón–Zygmund operator, we need to use the crucial estimate mentioned in Remark 2.3 to show that the kernel of the operator  $S_c(f)$  satisfies all conditions for the Calderón–Zygmund singular integral operator. We omit the details.

**Outline of proof of Theorem 4.4.** Since the proof (below) is rather complex, we begin by outlining our approach. For the first Plancherel–Pólya inequality (4.2), we substitute the wavelet reproducing formula (3.3) for  $f$  into the left-hand side. Thus the desired wavelet coefficients  $\langle \psi_\alpha^k, f \rangle$  appear. To deal with the unwanted terms  $D_{k'}$  and  $\psi_\alpha^k$ , we apply the almost-orthogonality estimates (4.4) given below. Then the standard technique, as in [DH], of applying an estimate from [FJ] and the Fefferman–Stein vector-valued maximal function inequality [FS] establishes (4.2).

The second Plancherel–Pólya inequality (4.3) is harder. Roughly speaking, we need to control the wavelet coefficients by the quantities  $D_{k'}(f)$ . Now for spaces of homogeneous type with additional assumptions, one proceeds as in [DH] via a frame reproducing formula of the form

$$f(x) = \sum_{k'} \sum_{\alpha' \in \mathcal{X}^{k'+N}} \mu(Q_{\alpha'}^{k'+N}) \tilde{D}_{k'}(x, x_{\alpha'}^{k'+N}) D_{k'}(f)(x_{\alpha'}^{k'+N}).$$

However, for our spaces of homogeneous type with no additional assumptions on  $d$  and  $\mu$ , no such frame reproducing formula is available. A new idea is needed. We introduce a suitable operator  $T_N$ , show that  $T_N$  is bounded and that the  $L^p(X)$  norm of  $S(T_N^{-1}(f))$  is controlled by that of  $S(f)$  (Lemma 4.6 below), and rewrite the wavelet coefficient as  $\langle \psi_{\alpha}^k, T_N^{-1} T_N f \rangle$ . Pulling out the operator  $T_N^{-1}$  from the left-hand side of (4.3), we obtain expressions of the form  $\langle \psi_{\alpha}^k / \sqrt{\mu(Q_{\alpha}^k)}, T_N f \rangle$ . Because of the form of  $T_N$ , we can now apply the almost-orthogonality estimates (4.4) to these terms and complete the remainder of the proof of the second Plancherel–Polya inequality (4.3) by following the approach used for (4.2).

We now give the details.

*Proof of Theorem 4.4.* For each  $f \in (\mathring{G}(\beta, \gamma))'$ , by Theorem 3.4, the functions

$$f_n(x) = \sum_{|k| \leq n} \sum_{\alpha \in \mathcal{Y}^k} \langle f, \psi_{\alpha}^k \rangle \psi_{\alpha}^k(x)$$

belong to  $L^2(X)$  and converge to  $f$  in  $(\mathring{G}(\beta, \gamma))'$  as  $n \rightarrow \infty$ . Note that  $\langle f_n, \psi_{\alpha}^k \rangle = \langle f, \psi_{\alpha}^k \rangle$  for  $|k| \leq n$ , and  $\langle f_n, \psi_{\alpha}^k \rangle = 0$  for  $|k| > n$ . Thus,

$$\begin{aligned} & \sum_{|k'| \leq n} \sum_{\alpha' \in \mathcal{X}^{k'+N}} \left[ \sup_{z \in Q_{\alpha'}^{k'+N}} |D_{k'}(f)(z)|^2 \right] \chi_{Q_{\alpha'}^{k'+N}}(x) \\ &= \sum_{k'} \sum_{\alpha' \in \mathcal{X}^{k'+N}} \left[ \sup_{z \in Q_{\alpha'}^{k'+N}} |D_{k'}(f_n)(z)|^2 \right] \chi_{Q_{\alpha'}^{k'+N}}(x) \end{aligned}$$

and

$$\sum_{|k| \leq n} \sum_{\alpha \in \mathcal{Y}^k} |\langle \psi_{\alpha}^k, f \rangle \tilde{\chi}_{Q_{\alpha}^k}(\cdot)|^2 = \sum_k \sum_{\alpha \in \mathcal{Y}^k} |\langle \psi_{\alpha}^k, f_n \rangle \tilde{\chi}_{Q_{\alpha}^k}(\cdot)|^2.$$

Therefore it suffices to show the inequality (4.2) of Theorem 4.4 for  $f \in L^2(X)$ , and similarly for the inequality (4.3).

We first prove (4.2). Fix  $N \in \mathbb{N}$ . The idea is to apply an almost-orthogonality estimate ((4.4) below). First, for each  $f \in L^2(X)$ , by the wavelet expansion (Theorem 3.4),

$$f(x) = \sum_{k \in \mathbb{Z}} \sum_{\alpha \in \mathcal{Y}^k} \langle f, \psi_{\alpha}^k \rangle \psi_{\alpha}^k(x).$$

Thus for each  $z \in Q_{\alpha'}^{k'+N}$  we have

$$D_{k'}(f)(z) = \sum_{k \in \mathbb{Z}} \sum_{\alpha \in \mathcal{Y}^k} \mu(Q_{\alpha}^k) \left\langle f, \frac{\psi_{\alpha}^k}{\sqrt{\mu(Q_{\alpha}^k)}} \right\rangle \left\langle \frac{\psi_{\alpha}^k(\cdot)}{\sqrt{\mu(Q_{\alpha}^k)}}, D_{k'}(\cdot, z) \right\rangle.$$

**Claim:** (*Almost-orthogonality estimate*) We claim that  $\langle \psi_{\alpha}^k(\cdot) / \sqrt{\mu(Q_{\alpha}^k)}, D_{k'}(\cdot, z) \rangle$  satisfies the following almost-orthogonality estimate: There exists a constant  $C$  such that for each

positive integer  $N$ , each  $\gamma > 0$ , each point  $z \in Q_{\alpha'}^{k'+N}$  and each point  $x_{\alpha'}^{k'+N} \in Q_{\alpha'}^{k'+N}$ , we have

$$(4.4) \quad \left| \left\langle \frac{\psi_{\alpha}^k(\cdot)}{\sqrt{\mu(Q_{\alpha}^k)}}, D_{k'}(\cdot, z) \right\rangle \right| \leq C \delta^{|k-k'|\eta} \frac{1}{V_{\delta(k' \wedge k)}(x_{\alpha}^k) + V_{\delta(k' \wedge k)}(x_{\alpha'}^{k'+N}) + V(x_{\alpha}^k, x_{\alpha'}^{k'+N})} \left( \frac{\delta^{(k \wedge k')}}{\delta^{(k \wedge k')} + d(x_{\alpha}^k, x_{\alpha'}^{k'+N})} \right)^{\gamma}.$$

As usual,  $k \wedge k' = \min\{k, k'\}$  denotes the minimum of  $k$  and  $k'$ .

**Remark 4.5.** The key idea used below to prove the claim (4.4) is that both  $\psi_{\alpha}^k(x)/\sqrt{\mu(Q_{\alpha}^k)}$  and  $D_{k'}(\cdot, z)$  satisfy size conditions, Hölder regularity conditions, and cancellation, since as we have shown,  $\psi_{\alpha}^k(x)/\sqrt{\mu(Q_{\alpha}^k)}$  is a test function in  $\mathring{G}(\beta, \gamma)$  while  $D_{k'}(\cdot, z)$  satisfies the properties (3.9)–(3.11) in Lemma 3.6. Further, we point out that if  $D_k(x, y)$  satisfies the same size condition (3.9) together with the following Hölder regularity condition (which is weaker than (3.10)),

$$(4.5) \quad |D_k(x, y) - D_k(x, y')| \leq C \left( \frac{d(y, y')}{\delta^k} \right)^{\eta} \left[ \frac{1}{V_{\delta^k}(x) + V(x, y)} \left( \frac{\delta^k}{\delta^k + d(x, y)} \right)^{\gamma} + \frac{1}{V_{\delta^k}(x) + V(x, y')} \left( \frac{\delta^k}{\delta^k + d(x, y')} \right)^{\gamma} \right],$$

and if the above estimate holds with  $x$  and  $y$  interchanged, then the above almost-orthogonality estimate (4.4) still holds, but with  $\eta$  replaced by some  $\eta' \in (0, \eta)$ .

Assuming the claim for the moment, we obtain that

$$\begin{aligned} & \sup_{z \in Q_{\alpha'}^{k'+N}} |D_{k'}(f)(z)| \\ & \leq C \sum_{k \in \mathbb{Z}} \sum_{\alpha \in \mathcal{Y}^k} \mu(Q_{\alpha}^k) \left| \left\langle f, \frac{\psi_{\alpha}^k}{\sqrt{\mu(Q_{\alpha}^k)}} \right\rangle \right| \delta^{|k-k'|\eta} \\ & \quad \times \frac{1}{V_{\delta(k' \wedge k)}(x_{\alpha}^k) + V_{\delta(k' \wedge k)}(x_{\alpha'}^{k'+N}) + V(x_{\alpha}^k, x_{\alpha'}^{k'+N})} \left( \frac{\delta^{(k \wedge k')}}{\delta^{(k \wedge k')} + d(x_{\alpha}^k, x_{\alpha'}^{k'+N})} \right)^{\gamma}. \end{aligned}$$

As a consequence, we have

$$\begin{aligned} & \left\{ \sum_{k'} \sum_{\alpha' \in \mathcal{Y}^{k'+N}} \sup_{z \in Q_{\alpha'}^{k'+N}} |D_{k'}(f)(z)|^2 \chi_{Q_{\alpha'}^{k'+N}}(x) \right\}^{1/2} \\ & \leq C \left\{ \sum_{k'} \sum_{\alpha' \in \mathcal{Y}^{k'+N}} \left| \sum_{k \in \mathbb{Z}} \sum_{\alpha \in \mathcal{Y}^k} \mu(Q_{\alpha}^k) \left\langle f, \frac{\psi_{\alpha}^k}{\sqrt{\mu(Q_{\alpha}^k)}} \right\rangle \right| \delta^{|k-k'|\eta} \right. \\ & \quad \times \left. \frac{1}{V_{\delta(k' \wedge k)}(x_{\alpha}^k) + V_{\delta(k' \wedge k)}(x) + V(x_{\alpha}^k, x)} \left( \frac{\delta^{(k \wedge k')}}{\delta^{(k \wedge k')} + d(x_{\alpha}^k, x)} \right)^{\gamma} \right\}^2 \chi_{Q_{\alpha'}^{k'+N}}(x) \Big\}^{1/2}. \end{aligned}$$

Using the same estimate as in [FJ], pp.147–148 (see also Lemma 2.12 in [HLL2]), we obtain

$$\begin{aligned} & \sum_{\alpha \in \mathcal{Y}^k} \mu(Q_{\alpha}^k) \frac{1}{V_{\delta(k' \wedge k)}(x_{\alpha}^k) + V_{\delta(k' \wedge k)}(x) + V(x_{\alpha}^k, x)} \left( \frac{\delta^{(k \wedge k')}}{\delta^{(k \wedge k')} + d(x_{\alpha}^k, x)} \right)^{\gamma} \left| \left\langle f, \frac{\psi_{\alpha}^k}{\sqrt{\mu(Q_{\alpha}^k)}} \right\rangle \right| \\ & \leq C \delta^{[(k \wedge k') - k]\omega(1-1/r)} \left\{ \mathcal{M} \left( \sum_{\alpha \in \mathcal{Y}^k} \left| \left\langle f, \frac{\psi_{\alpha}^k}{\sqrt{\mu(Q_{\alpha}^k)}} \right\rangle \right|^r \chi_{Q_{\alpha}^k}(\cdot) \right)(x) \right\}^{1/r}, \end{aligned}$$

where  $\mathcal{M}$  is the Hardy–Littlewood maximal function on  $X$  and  $\frac{\omega}{\omega+\eta} < r < p$ .

Thus, by the Fefferman–Stein vector-valued maximal function inequality with  $p/r > 1$  (see [FS]), we obtain

$$\begin{aligned} & \left\| \left\{ \sum_{k'} \sum_{\alpha' \in \mathcal{X}^{k'+N}} \sup_{z \in Q_{\alpha'}^{k'+N}} |D_{k'}(f)(z)| \chi_{Q_{\alpha'}^{k'+N}}(\cdot) \right\} \right\|_{L^p(X)}^{1/2} \\ & \leq C \left\| \left\{ \sum_k \sum_{\alpha \in \mathcal{Y}^k} |\langle \psi_\alpha^k, f \rangle \tilde{\chi}_{Q_\alpha^k}(\cdot)|^2 \right\} \right\|_{L^p(X)}^{1/2}. \end{aligned}$$

It remains to show the claimed almost-orthogonality estimate (4.4). We first consider the case  $k \geq k'$ . Applying the cancellation property for  $\psi_\alpha^k(x)$  yields

$$\begin{aligned} \left| \left\langle \frac{\psi_\alpha^k(\cdot)}{\sqrt{\mu(Q_\alpha^k)}}, D_{k'}(\cdot, z) \right\rangle \right| &= \left| \int_X \frac{\psi_\alpha^k(x)}{\sqrt{\mu(Q_\alpha^k)}} [D_{k'}(x, z) - D_{k'}(x_\alpha^k, z)] d\mu(x) \right| \\ &\leq \int_{W_1} \frac{|\psi_\alpha^k(x)|}{\sqrt{\mu(Q_\alpha^k)}} |D_{k'}(x, z) - D_{k'}(x_\alpha^k, z)| d\mu(x) \\ &\quad + \int_{W_2} \frac{|\psi_\alpha^k(x)|}{\sqrt{\mu(Q_\alpha^k)}} [|D_{k'}(x, z)| + |D_{k'}(x_\alpha^k, z)|] d\mu(x) \\ &=: U_1 + U_2, \end{aligned}$$

where  $W_1 := \{x \in X : d(x, x_\alpha^k) \leq (2A_0)^{-1}(\delta^{k'} + d(x_\alpha^k, z))\}$  and  $W_2 := X \setminus W_1$ .

Similarly to the estimate of  $(A)_1$  in the proof of Lemma 3.6, for  $U_1$ , using the size condition (Definition 3.1(i)) on  $\psi_\alpha^k(x)/\sqrt{\mu(Q_\alpha^k)}$  and the smoothness condition (Lemma 3.6(ii)) on  $D_{k'}(x, y)$ , we obtain that for all  $z, x_{\alpha'}^{k'+N} \in Q_{\alpha'}^{k'+N}$ ,

$$\begin{aligned} U_1 &\leq C \int_{W_1} \frac{1}{V_{\delta^k}(x_\alpha^k) + V(x_\alpha^k, x)} \left( \frac{\delta^k}{\delta^k + d(x_\alpha^k, x)} \right)^\Gamma \\ &\quad \times \left( \frac{d(x, x_\alpha^k)}{\delta^{k'} + d(x_\alpha^k, x_{\alpha'}^{k'+N})} \right)^\eta \frac{1}{V_{\delta^{k'}}(x_\alpha^k) + V(x, x_{\alpha'}^{k'+N})} \left( \frac{\delta^{k'}}{\delta^{k'} + d(x_\alpha^k, x_{\alpha'}^{k'+N})} \right)^\gamma d\mu(x) \\ &\leq C \delta^{(k-k')\eta} \int_{W_1} \frac{1}{V_{\delta^k}(x_\alpha^k) + V(x_\alpha^k, x)} \left( \frac{\delta^k}{\delta^k + d(x_\alpha^k, x)} \right)^{\Gamma-\eta} d\mu(x) \\ &\quad \times \frac{1}{V_{\delta^{k'}}(x_\alpha^k) + V_{\delta^{k'}}(x_{\alpha'}^{k'+N}) + V(x_\alpha^k, x_{\alpha'}^{k'+N})} \left( \frac{\delta^{k'}}{\delta^{k'} + d(x_\alpha^k, x_{\alpha'}^{k'+N})} \right)^\gamma \end{aligned}$$

for  $\Gamma > \eta$  and  $\gamma > 0$ .

The estimate for  $U_2$  is similar to the proof for  $(A)_2$  as in Lemma 3.6. Specifically, we have

$$\begin{aligned} U_2 &\leq C \int_{W_2} \frac{1}{V_{\delta^k}(x_\alpha^k) + V(x_\alpha^k, x)} \left( \frac{\delta^k}{\delta^k + d(x_\alpha^k, x)} \right)^\Gamma \left[ \frac{1}{V_{\delta^{k'}}(x) + V(x, x_{\alpha'}^{k'+N})} \left( \frac{\delta^{k'}}{\delta^{k'} + d(x, x_{\alpha'}^{k'+N})} \right)^\gamma \right. \\ &\quad \left. + \frac{1}{V_{\delta^{k'}}(x_\alpha^k) + V(x_\alpha^k, x_{\alpha'}^{k'+N})} \left( \frac{\delta^{k'}}{\delta^{k'} + d(x_\alpha^k, x_{\alpha'}^{k'+N})} \right)^\gamma \right] d\mu(x) \\ &\leq C \delta^{(k-k')\eta} \frac{1}{V_{\delta^{k'}}(x_\alpha^k) + V_{\delta^{k'}}(x_{\alpha'}^{k'+N}) + V(x_\alpha^k, x_{\alpha'}^{k'+N})} \left( \frac{\delta^{k'}}{\delta^{k'} + d(x_\alpha^k, x_{\alpha'}^{k'+N})} \right)^\gamma. \end{aligned}$$

These estimates of  $U_1$  and  $U_2$  establish the claimed almost-orthogonality estimate (4.4) when  $k \geq k'$ . The proof for the case  $k < k'$  is similar. This completes the proof of the



almost-orthogonality estimate (4.4), and hence the proof of the first Plancherel–Pólya inequality (4.2).

To show the second Plancherel–Pólya inequality (4.3), we need the following result about the operator  $T_N$ , as mentioned in the outline of the proof of Theorem 4.4.

**Lemma 4.6.** (Properties of  $T_N$ ) *Suppose that  $f \in L^2(X)$  and  $\frac{\omega}{\omega+\eta} < p < \infty$ , where  $\omega$  is the upper dimension of  $(X, d, \mu)$ . Let  $N$  be a positive integer. In each cube  $Q_\alpha^{k+N}$ , fix a point  $x_\alpha^{k+N}$ . Define the operator  $T_N$  by*

$$(4.6) \quad T_N(f)(x) := \sum_k \sum_{\alpha \in \mathcal{J}^{k+N}} \mu(Q_\alpha^{k+N}) D_k(x, x_\alpha^{k+N}) D_k(f)(x_\alpha^{k+N}).$$

Then the following assertions hold.

- (i)  $T_N$  is bounded on  $L^2(X)$ .
- (ii) There exists a constant  $C$  independent of  $f$  and of the choice of  $x_\alpha^{k+N}$  such that

$$\|S(T_N(f))\|_{L^p(X)} \leq C \|S(f)\|_{L^p(X)},$$

where  $S$  is the discrete Littlewood–Paley square function as in Definition 4.1.

- (iii) If  $N$  is chosen sufficiently large, then  $T_N$  is invertible and there is a constant  $C$  independent of  $f$  and of the choice of  $x_\alpha^{k+N}$  such that

$$(4.7) \quad \|S(T_N^{-1}(f))\|_{L^p(X)} \leq C \|S(f)\|_{L^p(X)}.$$

We defer the proof of this technical lemma until after the end of the proof of Theorem 4.4. We now continue the proof of the second Plancherel–Pólya inequality (4.3). Choose  $N$  sufficiently large that  $T_N$  is invertible and (4.7) holds. For  $f \in L^2(X)$ , write  $f = T_N^{-1} T_N f$ . Applying Lemma 4.6, we find that

$$(4.8) \quad \begin{aligned} \left\| \left\{ \sum_k \sum_{\alpha \in \mathcal{J}^k} |\langle \psi_\alpha^k, f \rangle \tilde{\chi}_{Q_\alpha^k}(\cdot)|^2 \right\}^{1/2} \right\|_{L^p(X)} &= \left\| \left\{ \sum_k \sum_{\alpha \in \mathcal{J}^k} |\langle \psi_\alpha^k, T_N^{-1} T_N f \rangle \tilde{\chi}_{Q_\alpha^k}(\cdot)|^2 \right\}^{1/2} \right\|_{L^p(X)} \\ &\leq C \left\| \left\{ \sum_k \sum_{\alpha \in \mathcal{J}^k} |\langle \psi_\alpha^k, T_N f \rangle \tilde{\chi}_{Q_\alpha^k}(\cdot)|^2 \right\}^{1/2} \right\|_{L^p(X)}. \end{aligned}$$

By the definition of  $T_N(f)$ , we have

$$\left\langle \frac{\psi_\alpha^k}{\sqrt{\mu(Q_\alpha^k)}}, T_N f \right\rangle = \sum_{k'} \sum_{\alpha' \in \mathcal{J}^{k'+N}} \mu(Q_{\alpha'}^{k'+N}) \left\langle \frac{\psi_\alpha^k(\cdot)}{\mu(Q_\alpha^k)}, D_{k'}(\cdot, x_{\alpha'}^{k'+N}) \right\rangle D_{k'}(f)(x_{\alpha'}^{k'+N}).$$

Therefore, for each fixed  $\eta' \in (0, \eta)$ ,

$$\begin{aligned} &\left\{ \sum_k \sum_{\alpha \in \mathcal{J}^k} |\langle \psi_\alpha^k, T_N f \rangle \tilde{\chi}_{Q_\alpha^k}(x)|^2 \right\}^{1/2} \\ &\leq C \left\{ \sum_k \sum_{\alpha \in \mathcal{J}^k} \left[ \sum_{k' \in \mathbb{Z}} \sum_{\alpha' \in \mathcal{J}^{k'+N}} \mu(Q_{\alpha'}^{k'+N}) \delta^{|k-k'|\eta'} \frac{1}{V_{\delta^{k'}}(x_\alpha^k) + V_{\delta^{k'}}(x_{\alpha'}^{k'+N}) + V(x_\alpha^k, x_{\alpha'}^{k'+N})} \right. \right. \\ &\quad \left. \left. \times \left( \frac{\delta^{(k \wedge k')}}{\delta^{(k \wedge k')} + d(x_\alpha^k, x_{\alpha'}^{k'+N})} \right)^\gamma D_{k'}(f)(x_{\alpha'}^{k'+N}) \chi_{Q_\alpha^k}(x) \right]^2 \right\}^{1/2} \\ &\leq C \left\{ \sum_k \sum_{\alpha \in \mathcal{J}^k} \left[ \sum_{k' \in \mathbb{Z}} \delta^{|k-k'|\eta'} \sum_{\alpha' \in \mathcal{J}^{k'+N}} \mu(Q_{\alpha'}^{k'+N}) \frac{1}{V_{\delta^{k'}}(x) + V_{\delta^{k'}}(x_{\alpha'}^{k'+N}) + V(x, x_{\alpha'}^{k'+N})} \right. \right. \end{aligned}$$

$$\times \left( \frac{\delta^{(k \wedge k')}}{\delta^{(k \wedge k')} + d(x, x_{\alpha'}^{k'+N})} \right)^\gamma D_{k'}(f)(x_{\alpha'}^{k'+N}) \chi_{Q_\alpha^k}(x) \Big]^2 \Big\}^{1/2}.$$

By the same estimate from [FJ] as in the proof of (4.2) above, we have

$$\begin{aligned} & \sum_{\alpha' \in \mathcal{X}^{k'+N}} \mu(Q_{\alpha'}^{k'+N}) \frac{1}{V_{\delta^{k'}}(x) + V_{\delta^{k'}}(x_{\alpha'}^{k'+N}) + V(x, x_{\alpha'}^{k'+N})} \\ & \times \left( \frac{\delta^{(k \wedge k')}}{\delta^{(k \wedge k')} + d(x, x_{\alpha'}^{k'+N})} \right)^\gamma D_{k'}(f)(x_{\alpha'}^{k'+N}) \\ & \leq C \delta^{[(k \wedge k') - k]\omega(1-1/r)} \left\{ \mathcal{M} \left( \sum_{\alpha' \in \mathcal{X}^{k'+N}} |D_{k'}(f)(x_{\alpha'}^{k'+N})|^r \chi_{Q_{\alpha'}^{k'+N}}(\cdot) \right)(x) \right\}^{1/r}, \end{aligned}$$

where  $\mathcal{M}$  is the Hardy–Littlewood maximal function on  $X$  and  $\frac{\omega}{\omega+\eta} < r < p$ . Note that the above inequality still holds when the point  $x_{\alpha'}^{k'+N}$  on the right-hand side is replaced by an arbitrary point  $z$  in  $Q_{\alpha'}^{k'+N}$ , and therefore also holds when the expression  $|D_{k'}(f)(x_{\alpha'}^{k'+N})|^r$  on the right-hand side is replaced by the infimum of  $|D_{k'}(f)(z)|^r$  over all  $z \in Q_{\alpha'}^{k'+N}$ . Thus, we have

$$\begin{aligned} & \left\{ \sum_k \sum_{\alpha \in \mathcal{Y}^k} |\langle \psi_\alpha^k, f \rangle \tilde{\chi}_{Q_\alpha^k}(x)|^2 \right\}^{1/2} \\ & \leq C \left\{ \sum_k \sum_{\alpha \in \mathcal{Y}^k} \left[ \sum_{k' \in \mathbb{Z}} \delta^{|k-k'| \epsilon} \delta^{[(k \wedge k') - k]\omega(1-1/r)} \right. \right. \\ & \quad \times \left. \left. \left\{ \mathcal{M} \left( \sum_{\alpha' \in \mathcal{X}^{k'+N}} \inf_{z \in Q_{\alpha'}^{k'+N}} |D_{k'}(f)(z)|^r \chi_{Q_{\alpha'}^{k'+N}}(\cdot) \right)(x) \right\}^{1/r} \right]^2 \chi_{Q_\alpha^k}(x) \right\}^{1/2}. \end{aligned}$$

Applying the Fefferman–Stein vector-valued maximal function inequality with  $p/r > 1$ , from [FS], we obtain

$$\begin{aligned} & \left\| \left\{ \sum_k \sum_{\alpha \in \mathcal{Y}^k} |\langle \psi_\alpha^k, f \rangle \tilde{\chi}_{Q_\alpha^k}(x)|^2 \right\}^{1/2} \right\|_{L^p(X)} \leq C \left\| \left\{ \sum_k \sum_{\alpha \in \mathcal{Y}^k} |\langle \psi_\alpha^k, T_N f \rangle \tilde{\chi}_{Q_\alpha^k}(x)|^2 \right\}^{1/2} \right\|_{L^p(X)} \\ & \leq C \left\| \left\{ \sum_{k' \in \mathbb{Z}} \left[ \mathcal{M} \left( \sum_{\alpha' \in \mathcal{X}^{k'+N}} \inf_{z \in Q_{\alpha'}^{k'+N}} |D_{k'}(f)(z)|^r \chi_{Q_{\alpha'}^{k'+N}}(\cdot) \right)(x) \right]^{2/r} \right\}^{1/2} \right\|_{L^p(X)} \\ & \leq C \left\| \left\{ \sum_{k'} \sum_{\alpha' \in \mathcal{X}^{k'+N}} \inf_{z \in Q_{\alpha'}^{k'+N}} |D_{k'}(f)(z)| \chi_{Q_{\alpha'}^{k'+N}}(x) \right\}^{1/2} \right\|_{L^p(X)}, \end{aligned}$$

which implies that the second Plancherel–Pólya inequality (4.3) holds for  $f \in L^2(X)$ . The proof of Theorem 4.4 is complete, except for the proof of Lemma 4.6.  $\square$

It remains to prove the technical lemma used in the preceding proof.

*Proof of Lemma 4.6.* (i) Fix  $N \in \mathbb{N}$ . We show that the operator  $T_N$  is bounded on  $L^2(X)$ . Write  $T_N(f)(x) = \sum_k E_k(f)(x)$ , where the kernel  $E_k(x, y)$  of  $E_k$  is given by

$$E_k(x, y) := \sum_{\alpha \in \mathcal{X}^{k+N}} \mu(Q_\alpha^{k+N}) D_k(x, x_\alpha^{k+N}) D_k(x_\alpha^{k+N}, y).$$

This kernel  $E_k(x, y)$  satisfies the same decay and smoothness estimates (3.9) and (3.10) as  $D_k(x, y)$  does, with bounds independent of  $x_\alpha^{k+N}$ , as can be shown by a proof similar to that for  $D_k(x, y)$ . Moreover,  $\int_X E_k(x, y) d\mu(y) = 0$  for each  $x \in X$  and  $\int_X E_k(x, y) d\mu(x) = 0$  for

each  $y \in X$ . Therefore the Cotlar–Stein lemma can be applied to show that  $T_N$  is bounded on  $L^2(X)$ .

(ii) Suppose that  $f \in L^2(X)$  and  $\frac{\omega}{\omega+\eta} < p < \infty$ . Then by the definition of  $D_k$  and the wavelet reproducing formula (2.13), we have

$$\begin{aligned} f(x) &= \sum_k D_k D_k(f)(x) = \sum_k \sum_{\alpha \in \mathcal{X}^{k+N}} \mu(Q_\alpha^{k+N}) D_k(x, x_\alpha^{k+N}) D_k(f)(x_\alpha^{k+N}) \\ &\quad + \left( \sum_k D_k D_k(f)(x) - \sum_k \sum_{\alpha \in \mathcal{X}^{k+N}} \mu(Q_\alpha^{k+N}) D_k(x, x_\alpha^{k+N}) D_k(f)(x_\alpha^{k+N}) \right) \\ &=: T_N(f)(x) + R_N(f)(x), \end{aligned}$$

where  $x_\alpha^{k+N}$  are arbitrary fixed points in  $Q_\alpha^{k+N}$ .

Since  $T_N = I - R_N$  by definition, to show (ii) in Lemma 4.6, it suffices to show that

$$(4.9) \quad \|S(R_N(f))\|_{L^p(X)} \leq C\delta^{\eta N} \|S(f)\|_{L^p(X)}.$$

For then

$$\|S(T_N(f))\|_{L^p(X)} \leq \|S(f)\|_{L^p(X)} + \|S(R_N(f))\|_{L^p(X)} \leq (1 + C\delta^{\eta N}) \|S(f)\|_{L^p(X)},$$

as required.

To establish (4.9), we write

$$R_N(f)(x) = \sum_k \sum_{\alpha \in \mathcal{X}^{k+N}} \int_{Q_\alpha^{k+N}} [D_k(x, z) D_k(f)(z) - D_k(x, x_\alpha^{k+N}) D_k(f)(x_\alpha^{k+N})] d\mu(z).$$

Thus the kernel  $R_N(x, y)$  of  $R_N$  is given by

$$\begin{aligned} R_N(x, y) &:= \sum_k \sum_{\alpha \in \mathcal{X}^{k+N}} \int_{Q_\alpha^{k+N}} [D_k(x, z) D_k(z, y) - D_k(x, x_\alpha^{k+N}) D_k(x_\alpha^{k+N}, y)] d\mu(z) \\ &= \sum_k \sum_{\alpha \in \mathcal{X}^{k+N}} \int_{Q_\alpha^{k+N}} [D_k(x, z) - D_k(x, x_\alpha^{k+N})] D_k(z, y) d\mu(z) \\ &\quad + \sum_k \sum_{\alpha \in \mathcal{X}^{k+N}} \int_{Q_\alpha^{k+N}} D_k(x, x_\alpha^{k+N}) [D_k(z, y) - D_k(x_\alpha^{k+N}, y)] d\mu(z) \\ &=: R_N^{(1)}(x, y) + R_N^{(2)}(x, y). \end{aligned}$$

Note that by the same proof as for  $T_N$ , both  $R_N^{(1)}$  and  $R_N^{(2)}$  are bounded on  $L^2(X)$ , and therefore the inner products  $\langle \psi_\alpha^k, R_N^{(1)}(f) \rangle$  and  $\langle \psi_\alpha^k, R_N^{(2)}(f) \rangle$  are well defined. To estimate  $\|S(R_N^{(1)}(f))\|_{L^p(X)}$ , we write

$$\|S(R_N^{(1)}(f))\|_{L^p(X)} = \left\| \left\{ \sum_k \sum_{\alpha \in \mathcal{X}^k} |\langle \psi_\alpha^k, R_N^{(1)}(f) \rangle \tilde{\chi}_{Q_\alpha^k}(\cdot)|^2 \right\}^{1/2} \right\|_{L^p(X)}.$$

By the  $L^2(X)$ -boundedness of  $R_N^{(1)}$  and the wavelet reproducing formula (2.13) for  $f \in L^2(X)$ , we have

$$\begin{aligned} &\left\langle \frac{\psi_\alpha^k}{\sqrt{\mu(Q_\alpha^k)}}, R_N^{(1)}(f) \right\rangle \\ &= \int_X \frac{\psi_\alpha^k(x)}{\sqrt{\mu(Q_\alpha^k)}} \sum_{k'} \sum_{\alpha' \in \mathcal{X}^{k'+N}} \int_{Q_{\alpha'}^{k'+N}} [D_{k'}(x, z) - D_{k'}(x, x_{\alpha'}^{k'+N})] \end{aligned}$$

$$\begin{aligned}
& \times \int_X D_{k'}(z, y) \sum_{k'' \in \mathbb{Z}} \sum_{\alpha'' \in \mathcal{Y}^{k''}} \langle \psi_{\alpha''}^{k''}, f \rangle \psi_{\alpha''}^{k''}(y) d\mu(y) d\mu(z) d\mu(x) \\
&= \sum_{k''} \sum_{\alpha'' \in \mathcal{Y}^{k''}} \mu(Q_{\alpha''}^{k''}) \left\langle \frac{\psi_{\alpha''}^{k''}}{\sqrt{\mu(Q_{\alpha''}^{k''})}}, f \right\rangle \sum_{k'} \int_X \int_X \frac{\psi_{\alpha}^k(x)}{\sqrt{\mu(Q_{\alpha}^k)}} \overline{D}_{k'}(x, y) \frac{\psi_{\alpha''}^{k''}(y)}{\sqrt{\mu(Q_{\alpha''}^{k''})}} d\mu(y) d\mu(x) \\
&= \sum_{k''} \sum_{\alpha'' \in \mathcal{Y}^{k''}} \mu(Q_{\alpha''}^{k''}) \left\langle \frac{\psi_{\alpha''}^{k''}}{\sqrt{\mu(Q_{\alpha''}^{k''})}}, f \right\rangle \left\langle \frac{\psi_{\alpha}^k(\cdot)}{\sqrt{\mu(Q_{\alpha}^k)}}, F_{k''}(\cdot, x_{\alpha''}^{k''}) \right\rangle,
\end{aligned}$$

where

$$\overline{D}_{k'}(x, y) := \sum_{\alpha' \in \mathcal{X}^{k'+N}} \int_{Q_{\alpha'}^{k'+N}} [D_{k'}(x, z) - D_{k'}(x, x_{\alpha'}^{k'+N})] D_{k'}(z, y) d\mu(z)$$

and

$$F_{k''}(x, x_{\alpha''}^{k''}) := \sum_{k'} \int_X \overline{D}_{k'}(x, y) \frac{\psi_{\alpha''}^{k''}(y)}{\sqrt{\mu(Q_{\alpha''}^{k''})}} d\mu(y).$$

We now show that  $\left\langle \psi_{\alpha}^k(\cdot)/\sqrt{\mu(Q_{\alpha}^k)}, F_{k''}(\cdot, x_{\alpha''}^{k''}) \right\rangle$  satisfies an almost-orthogonality estimate similar to (4.4), by following the philosophy of Remark 4.5. Recall that  $\psi_{\alpha}^k(\cdot)/\sqrt{\mu(Q_{\alpha}^k)}$  is a test function. It remains to show that the function  $F_{k''}(x, x_{\alpha''}^{k''})$  satisfies a size condition, a Hölder regularity condition, and cancellation.

Next, it seems unlikely that  $F_{k''}(x, x_{\alpha''}^{k''})$  satisfies the Hölder regularity condition (3.10). However, as noted in Remark 4.5, it suffices to establish the weaker Hölder regularity condition (4.5), which we now do. To begin, we show that  $F_{k''}(x, x_{\alpha''}^{k''})$  satisfies

$$(a)' \quad |F_{k''}(x, x_{\alpha''}^{k''})| \leq C\delta^{\eta N} \frac{1}{V_{\delta^{k''}}(x) + V(x, x_{\alpha''}^{k''})} \left( \frac{\delta^{k''}}{\delta^{k''} + d(x, x_{\alpha''}^{k''})} \right)^{\gamma},$$

for all  $\gamma \in (0, \eta)$ , and

$$\begin{aligned}
(b)' \quad & |F_{k''}(x, x_{\alpha''}^{k''}) - F_{k''}(x', x_{\alpha''}^{k''})| \\
& \leq C\delta^{\eta N} \left( \frac{d(x, x')}{\delta^{k'}} \right)^{\eta'} \times \left[ \frac{1}{V_{\delta^{k''}}(x) + V(x, x_{\alpha''}^{k''})} \left( \frac{\delta^{k''}}{\delta^{k''} + d(x, x_{\alpha''}^{k''})} \right)^{\gamma} \right. \\
& \quad \left. + \frac{1}{V_{\delta^{k''}}(x') + V(x', x_{\alpha''}^{k''})} \left( \frac{\delta^{k''}}{\delta^{k''} + d(x', x_{\alpha''}^{k''})} \right)^{\gamma} \right],
\end{aligned}$$

for all  $\eta' \in (0, \eta)$ .

To prove (a)' and (b)', we first show that  $\overline{D}_{k'}(x, y)$  satisfies the same estimates (3.9) and (3.10) as  $D_k(x, y)$ , but with the constant  $C$  replaced by  $C\delta^{\eta N}$ , that is,

$$(a) \quad |\overline{D}_{k'}(x, y)| \leq C\delta^{\eta N} \frac{1}{V_{\delta^{k'}}(x) + V(x, y)} \left( \frac{\delta^{k'}}{\delta^{k'} + d(x, y)} \right)^{\gamma};$$

$$(b) \quad |\overline{D}_{k'}(x, y) - \overline{D}_{k'}(x', y)| \leq C\delta^{\eta N} \left( \frac{d(x, x')}{\delta^{k'}} \right)^{\eta} \frac{1}{V_{\delta^{k'}}(x) + V(x, y)} \left( \frac{\delta^{k'}}{\delta^{k'} + d(x, y)} \right)^{\gamma}$$

for  $d(x, x') \leq (2A_0)^{-1} \max\{\delta^{k'} + d(x, y), \delta^{k'} + d(x', y)\}$ , and

$$(c) \quad |\overline{D}_{k'}(x, y) - \overline{D}_{k'}(x, y')| \leq C\delta^{\eta N} \left( \frac{d(y, y')}{\delta^{k'} + d(x, y)} \right)^\eta \frac{1}{V_{\delta^{k'}}(x) + V(x, y)} \left( \frac{\delta^{k'}}{\delta^{k'} + d(x, y)} \right)^\gamma$$

for  $d(y, y') \leq (2A_0)^{-1} \max\{\delta^{k'} + d(x, y), \delta^{k'} + d(x, y')\}$ . Moreover,

$$\int_X \overline{D}_{k'}(x, y) d\mu(x) = 0 \quad \text{and} \quad \int_X \overline{D}_{k'}(x, y) d\mu(y) = 0,$$

for all  $y \in X$  and all  $x \in X$ , respectively. Indeed, note that  $[D_{k'}(x, z) - D_{k'}(x, x_{\alpha'}^{k'+N})]$  satisfies the same estimates (3.9) and (3.10) as  $D_{k'}(x, z)$  does, but with the constant  $C$  replaced by  $C\delta^{\eta N}$ . Therefore, the proofs for (a), (b) and (c) follow from a similar proof to that for Lemma 3.6. As a consequence, the almost-orthogonality estimate (4.4) holds for  $\langle \overline{D}_{k'}(x, \cdot), \psi_{\alpha''}^{k''}(\cdot) / \sqrt{\mu(Q_{\alpha''}^{k''})} \rangle$ . We omit the details.

Now, to verify the estimate in (a)', applying this almost-orthogonality estimate yields that

$$\begin{aligned} |F_{k''}(x, x_{\alpha''}^{k''})| &\leq \sum_{k'} \left| \left\langle \overline{D}_{k'}(x, \cdot), \frac{\psi_{\alpha''}^{k''}(\cdot)}{\sqrt{\mu(Q_{\alpha''}^{k''})}} \right\rangle \right| \\ &\leq C\delta^{\eta N} \sum_{k'} \delta^{|k'-k''|_\eta} \frac{1}{V_{\delta^{(k' \wedge k'')}}(x) + V(x, x_{\alpha''}^{k''})} \left( \frac{\delta^{(k' \wedge k'')}}{\delta^{(k' \wedge k'')} + d(x, x_{\alpha''}^{k''})} \right)^\gamma \\ &\leq C\delta^{\eta N} \sum_{k'} \delta^{|k-k'|(\eta-\gamma)} \frac{1}{V_{\delta^{k''}}(x) + V(x, x_{\alpha''}^{k''})} \left( \frac{\delta^{k''}}{\delta^{k''} + d(x, x_{\alpha''}^{k''})} \right)^\gamma \\ &\leq C\delta^{\eta N} \frac{1}{V_{\delta^{k''}}(x) + V(x, x_{\alpha''}^{k''})} \left( \frac{\delta^{k''}}{\delta^{k''} + d(x, x_{\alpha''}^{k''})} \right)^\gamma, \end{aligned}$$

for all  $\gamma \in (0, \eta)$ .

Next we show the estimate in (b)'. Note that

$$\begin{aligned} |\overline{D}_{k'}(x, y) - \overline{D}_{k'}(x', y)| &\leq C\delta^{\eta N} \left( \frac{d(x, x')}{\delta^{k'}} \right)^\eta \\ &\times \left[ \frac{1}{V_{\delta^{k'}}(x) + V(x, y)} \left( \frac{\delta^{k'}}{\delta^{k'} + d(x, y)} \right)^\gamma + \frac{1}{V_{\delta^{k'}}(x') + V(x', y)} \left( \frac{\delta^{k'}}{\delta^{k'} + d(x', y)} \right)^\gamma \right] \end{aligned}$$

and

$$\int_X [\overline{D}_{k'}(x, y) - \overline{D}_{k'}(x', y)] d\mu(y) = 0.$$

Therefore, as pointed out in Remark 4.5, we obtain for  $k' > k''$  that

$$\begin{aligned} |E_{k''}(x, x_{\alpha''}^{k''}) - E_{k''}(x', x_{\alpha''}^{k''})| &\leq \sum_{k'} \left| \left\langle [\overline{D}_{k'}(x, \cdot) - \overline{D}_{k'}(x', \cdot)], \frac{\psi_{\alpha''}^{k''}(\cdot)}{\sqrt{\mu(Q_{\alpha''}^{k''})}} \right\rangle \right| \\ &\leq C\delta^{\eta N} \sum_{k'} \delta^{|k'-k''|_\eta} \left( \frac{d(x, x')}{\delta^{k'}} \right)^{\eta'} \left[ \frac{1}{V_{\delta^{k''}}(x) + V(x, x_{\alpha''}^{k''})} \left( \frac{\delta^{k''}}{\delta^{k''} + d(x, x_{\alpha''}^{k''})} \right)^\gamma \right. \\ &\quad \left. + \frac{1}{V_{\delta^{k''}}(x') + V(x', x_{\alpha''}^{k''})} \left( \frac{\delta^{k''}}{\delta^{k''} + d(x', x_{\alpha''}^{k''})} \right)^\gamma \right] \end{aligned}$$

$$\begin{aligned}
&\leq C\delta^{\eta N} \sum_{k'} \delta^{|k'-k''|(\eta-\eta')} \left( \frac{d(x, x')}{\delta^{k''}} \right)^{\eta'} \left[ \frac{1}{V_{\delta^{k''}}(x) + V(x, x_{\alpha''}^{k''})} \left( \frac{\delta^{k''}}{\delta^{k''} + d(x, x_{\alpha''}^{k''})} \right)^{\gamma} \right. \\
&\quad \left. + \frac{1}{V_{\delta^{k''}}(x') + V(x', x_{\alpha''}^{k''})} \left( \frac{\delta^{k''}}{\delta^{k''} + d(x', x_{\alpha''}^{k''})} \right)^{\gamma} \right], \\
&\leq C\delta^{\eta N} \left( \frac{d(x, x')}{\delta^{k''}} \right)^{\eta'} \left[ \frac{1}{V_{\delta^{k''}}(x) + V(x, x_{\alpha''}^{k''})} \left( \frac{\delta^{k''}}{\delta^{k''} + d(x, x_{\alpha''}^{k''})} \right)^{\gamma} \right. \\
&\quad \left. + \frac{1}{V_{\delta^{k''}}(x') + V(x', x_{\alpha''}^{k''})} \left( \frac{\delta^{k''}}{\delta^{k''} + d(x', x_{\alpha''}^{k''})} \right)^{\gamma} \right],
\end{aligned}$$

for all  $\eta' \in (0, \eta)$ .

For  $k' \leq k''$ , we have

$$\begin{aligned}
&|E_{k''}(x, x_{\alpha''}^{k''}) - E_{k''}(x', x_{\alpha''}^{k''})| \\
&\leq \sum_{k' < k''} \int_X |[\bar{D}_{k'}(x, y) - \bar{D}_{k'}(x', y)]| \left| \frac{\psi_{\alpha''}^{k''}(y)}{\sqrt{\mu(Q_{\alpha''}^{k''})}} \right| d\mu(y) \\
&\leq C\delta^{\eta N} \sum_{k' \leq k''} \int_X \left( \frac{d(x, x')}{\delta^{k'}} \right)^{\eta} \left[ \frac{1}{V_{\delta^{k'}}(x) + V(x, y)} \left( \frac{\delta^{k'}}{\delta^{k'} + d(x, y)} \right)^{\gamma} \right. \\
&\quad \left. + \frac{1}{V_{\delta^{k'}}(x') + V(x', y)} \left( \frac{\delta^{k'}}{\delta^{k'} + d(x', y)} \right)^{\gamma} \right] \frac{1}{V_{\delta^{k''}}(y) + V(y, x_{\alpha''}^{k''})} \left( \frac{\delta^{k''}}{\delta^{k''} + d(y, x_{\alpha''}^{k''})} \right)^{\gamma} d\mu(y) \\
&\leq C\delta^{\eta N} \left( \frac{d(x, x')}{\delta^{k''}} \right)^{\eta} \sum_{k' \leq k''} \delta^{\eta(k''-k')} \left[ \frac{1}{V_{\delta^{k'}}(x) + V(x, x_{\alpha''}^{k''})} \left( \frac{\delta^{k'}}{\delta^{k'} + d(x, x_{\alpha''}^{k''})} \right)^{\gamma} \right. \\
&\quad \left. + \frac{1}{V_{\delta^{k'}}(x') + V(x', x_{\alpha''}^{k''})} \left( \frac{\delta^{k'}}{\delta^{k'} + d(x', x_{\alpha''}^{k''})} \right)^{\gamma} \right] \\
&\leq C\delta^{\eta N} \left( \frac{d(x, x')}{\delta^{k''}} \right)^{\eta} \sum_{k' \leq k''} \delta^{(\eta-\gamma)(k''-k')} \left[ \frac{1}{V_{\delta^{k''}}(x) + V(x, x_{\alpha''}^{k''})} \left( \frac{\delta^{k''}}{\delta^{k''} + d(x, x_{\alpha''}^{k''})} \right)^{\gamma} \right. \\
&\quad \left. + \frac{1}{V_{\delta^{k''}}(x') + V(x', x_{\alpha''}^{k''})} \left( \frac{\delta^{k''}}{\delta^{k''} + d(x', x_{\alpha''}^{k''})} \right)^{\gamma} \right] \\
&\leq C\delta^{\eta N} \left( \frac{d(x, x')}{\delta^{k''}} \right)^{\eta} \left[ \frac{1}{V_{\delta^{k''}}(x) + V(x, x_{\alpha''}^{k''})} \left( \frac{\delta^{k''}}{\delta^{k''} + d(x, x_{\alpha''}^{k''})} \right)^{\gamma} \right. \\
&\quad \left. + \frac{1}{V_{\delta^{k''}}(x') + V(x', x_{\alpha''}^{k''})} \left( \frac{\delta^{k''}}{\delta^{k''} + d(x', x_{\alpha''}^{k''})} \right)^{\gamma} \right],
\end{aligned}$$

for all  $\gamma \in (0, \eta)$ .

With the almost-orthogonality estimate in hand, the same argument as for (4.2), via the estimate from [FJ] and the Fefferman–Stein vector-valued maximal function, yields

$$\|S(R_N^{(1)}(f))\|_{L^p(X)} = \left\| \left\{ \sum_k \sum_{\alpha \in \mathcal{Y}^k} |\langle \psi_{\alpha}^k, R_N^{(1)}(f) \rangle \tilde{\chi}_{Q_{\alpha}^k}(\cdot)|^2 \right\}^{1/2} \right\|_{L^p(X)}$$

$$\leq C\delta^{\eta N} \left\| \left\{ \sum_k \sum_{\alpha \in \mathcal{G}^k} |\langle \psi_\alpha^k, (f) \rangle \tilde{\chi}_{Q_\alpha^k}(\cdot)|^2 \right\}^{1/2} \right\|_{L^p(X)} = C\delta^{\eta N} \|S(f)\|_{L^p(X)}.$$

A similar proof shows that  $\|S(R_N^{(2)}(f))\|_{L^p(X)} \leq C\delta^{\eta N} \|S(f)\|_{L^p(X)}$ . Therefore (4.9) holds:

$$\|S(R_N(f))\|_{L^p(X)} \leq C\delta^{\eta N} \|S(f)\|_{L^p(X)},$$

as required.

(iii) Consider the Neumann series  $(T_N)^{-1} = (I - R_N)^{-1} = \sum_{i=0}^{\infty} (R_N)^i$ . By (4.9) we have

$$\|S((T_N)^{-1}(f))\|_{L^p(X)} \leq \sum_{i=0}^{\infty} \|S((R_N)^i(f))\|_{L^p(X)} \leq (1 - C\delta^{\eta N})^{-1} \|S(f)\|_{L^p(X)},$$

as required, if  $N$  is chosen sufficiently large that  $C\delta^{\eta N} < 1$ .

This completes the proof of Lemma 4.6.  $\square$

We turn to the product setting.

#### 4.2. Product square functions via wavelets, and Plancherel–Pólya inequalities.

We now assume that  $\tilde{X} = X_1 \times X_2$  where each  $X_i$  is a space of homogeneous type as above. In this subsection,  $(x_1, x_2)$  denotes an element of  $X_1 \times X_2$ .

**Definition 4.7.** (Product square functions) Take  $\beta_i \in (0, \eta_i)$  and  $\gamma_i > 0$ , for  $i = 1, 2$ , and consider  $f \in (\mathring{G}(\beta_1, \beta_2; \gamma_1, \gamma_2))'$ .

(a) The *discrete product Littlewood–Paley square function*  $\tilde{S}(f)$  in terms of wavelet coefficients is defined by

$$(4.10) \quad \tilde{S}(f)(x_1, x_2) := \left\{ \sum_{k_1} \sum_{\alpha_1 \in \mathcal{G}^{k_1}} \sum_{k_2} \sum_{\alpha_2 \in \mathcal{G}^{k_2}} \left| \langle \psi_{\alpha_1}^{k_1} \psi_{\alpha_2}^{k_2}, f \rangle \tilde{\chi}_{Q_{\alpha_1}^{k_1}}(x_1) \tilde{\chi}_{Q_{\alpha_2}^{k_2}}(x_2) \right|^2 \right\}^{1/2},$$

where  $\psi_{\alpha_1}^{k_1} \psi_{\alpha_2}^{k_2} = \psi_{\alpha_1}^{k_1} \otimes \psi_{\alpha_2}^{k_2}$  with  $\psi_{\alpha_i}^{k_i}$  acting on the  $X_i$  variable for  $i = 1, 2$ , and  $\tilde{\chi}_{Q_{\alpha_i}^{k_i}}(x_i) := \chi_{Q_{\alpha_i}^{k_i}}(x_i) \mu_i(Q_{\alpha_i}^{k_i})^{-1/2}$ .

(b) The *continuous product Littlewood–Paley square function*  $\tilde{S}_c(f)$  in terms of wavelet operators is defined by

$$(4.11) \quad \tilde{S}_c(f)(x_1, x_2) := \left\{ \sum_{k_1} \sum_{k_2} \left| D_{k_1} D_{k_2}(f)(x_1, x_2) \right|^2 \right\}^{1/2},$$

where  $D_{k_i} := \sum_{\alpha_i \in \mathcal{G}^{k_i}} \psi_{\alpha_i}^{k_i}$  for  $i = 1, 2$ , and  $D_{k_1} D_{k_2} := D_{k_1} \otimes D_{k_2}$ .

The main results of this subsection are the following product versions of the Littlewood–Paley theory and the Plancherel–Pólya inequalities.

**Theorem 4.8.** (Product Littlewood–Paley theory) Suppose  $\beta_i \in (0, \eta_i)$ ,  $\gamma_i > 0$ , and  $\max\{\frac{\omega_1}{\omega_1 + \eta_1}, \frac{\omega_2}{\omega_2 + \eta_2}\} < p < \infty$ , where  $\omega_i$  is the upper dimension of  $X_i$ , for  $i = 1, 2$ . For all  $f \in (\mathring{G}(\beta_1, \beta_2; \gamma_1, \gamma_2))'$ , we have

$$\|\tilde{S}(f)\|_{L^p(X_1 \times X_2)} \sim \|\tilde{S}_c(f)\|_{L^p(X_1 \times X_2)}.$$

Moreover, if  $1 < p < \infty$ , then

$$\|\tilde{S}(f)\|_{L^p(X_1 \times X_2)} \sim \|\tilde{S}_c(f)\|_{L^p(X_1 \times X_2)} \sim \|f\|_{L^p(X_1 \times X_2)}.$$

**Theorem 4.9.** (Product Plancherel–Pólya inequalities) *Suppose  $\beta_i \in (0, \eta_i)$ ,  $\gamma_i > 0$ , and  $\max\{\frac{\omega_1}{\omega_1+\eta_1}, \frac{\omega_2}{\omega_2+\eta_2}\} < p < \infty$ , where  $\omega_i$  is the upper dimension of  $X_i$ , for  $i = 1, 2$ . Take  $N_1, N_2 \in \mathbb{N}$ . Then there is a positive constant  $C$  such that for all  $f \in (\mathring{G}(\beta_1, \beta_2; \gamma_1, \gamma_2))'$ , we have*

$$(4.12) \quad \left\| \left\{ \sum_{k'_1} \sum_{\alpha'_1 \in \mathcal{X}^{k'_1+N_1}} \sum_{k'_2} \sum_{\alpha'_2 \in \mathcal{X}^{k'_2+N_2}} \sup_{(z_1, z_2) \in Q_{\alpha'_1}^{k'_1+N_1} \times Q_{\alpha'_2}^{k'_2+N_2}} |D_{k'_1} D_{k'_2}(f)(z_1, z_2)| \chi_{Q_{\alpha'_1}^{k'_1+N_1}}(\cdot) \chi_{Q_{\alpha'_2}^{k'_2+N_2}}(\cdot) \right\}^{1/2} \right\|_{L^p(X_1 \times X_2)} \\ \leq C \left\| \left\{ \sum_{k_1} \sum_{\alpha_1 \in \mathcal{Y}^{k_1}} \sum_{k_2} \sum_{\alpha_2 \in \mathcal{Y}^{k_2}} |\langle \psi_{\alpha_1}^{k_1} \psi_{\alpha_2}^{k_2}, f \rangle \tilde{\chi}_{Q_{\alpha_1}^{k_1}}(\cdot) \tilde{\chi}_{Q_{\alpha_2}^{k_2}}(\cdot)|^2 \right\}^{1/2} \right\|_{L^p(X_1 \times X_2)}.$$

Further, suppose  $N_1$  and  $N_2$  are sufficiently large positive integers, to be determined during the proof below. Then there is a positive constant  $C$  such that for all  $f \in (\mathring{G}(\beta_1, \beta_2; \gamma_1, \gamma_2))'$ , we have

$$(4.13) \quad \left\| \left\{ \sum_{k_1} \sum_{\alpha_1 \in \mathcal{Y}^{k_1}} \sum_{k_2} \sum_{\alpha_2 \in \mathcal{Y}^{k_2}} |\langle \psi_{\alpha_1}^{k_1} \psi_{\alpha_2}^{k_2}, f \rangle \tilde{\chi}_{Q_{\alpha_1}^{k_1}}(\cdot) \tilde{\chi}_{Q_{\alpha_2}^{k_2}}(\cdot)|^2 \right\}^{1/2} \right\|_{L^p(X_1 \times X_2)} \\ \leq C \left\| \left\{ \sum_{k'_1} \sum_{\alpha'_1 \in \mathcal{X}^{k'_1+N_1}} \sum_{k'_2} \sum_{\alpha'_2 \in \mathcal{X}^{k'_2+N_2}} \inf_{(z_1, z_2) \in Q_{\alpha'_1}^{k'_1+N_1} \times Q_{\alpha'_2}^{k'_2+N_2}} |D_{k'_1} D_{k'_2}(f)(z_1, z_2)| \chi_{Q_{\alpha'_1}^{k'_1+N_1}}(\cdot) \chi_{Q_{\alpha'_2}^{k'_2+N_2}}(\cdot) \right\}^{1/2} \right\|_{L^p(X_1 \times X_2)}.$$

*Proofs of Theorems 4.8 and 4.9.* The proofs of Theorems 4.8 and 4.9 are analogous to those for the case of one factor. As mentioned in that case, the proofs of Theorems 4.3 and 4.4 follow from the almost-orthogonality estimates, namely the claim (4.4). To see that these proofs can be carried over to the product case, we make an observation analogous to the claim (4.4), as follows:

$$\left\langle D_{k_1} D_{k_2}(x_1, x_2, \cdot, \cdot), \psi_{\alpha_1}^{k_1} \psi_{\alpha_2}^{k_2}(\cdot, \cdot, y_1, y_2) \right\rangle = \left\langle D_{k_1}(x_1, \cdot), \psi_{\alpha_1}^{k_1}(\cdot, y_1) \right\rangle \left\langle D_{k_2}(x_2, \cdot), \psi_{\alpha_2}^{k_2}(\cdot, y_2) \right\rangle,$$

which together with the almost-orthogonality estimates for the one-factor case yields the desired almost-orthogonality estimates for the product case. For the product case, all estimates analogous to those in (a)–(c), (a)'–(b)', (4.2) and (4.3) follow similarly. We omit the details.  $\square$

## 5. PRODUCT $H^p$ , $\text{CMO}^p$ , $\text{BMO}$ AND $\text{VMO}$ , AND DUALITY

In this section we define the Hardy spaces  $H^p$ , the Carleson measure spaces  $\text{CMO}^p$  (including the bounded mean oscillation space  $\text{BMO} = \text{CMO}^1$ ), and the vanishing mean oscillation space  $\text{VMO}$ , in the setting of product spaces of homogeneous type. Both  $H^p$  and  $\text{CMO}^p$  are defined here for  $p$  in the range  $\max\{\frac{\omega_1}{\omega_1+\eta_1}, \frac{\omega_2}{\omega_2+\eta_2}\} < p \leq 1$ , where  $\omega_i$  is the upper dimension of  $X_i$ , for  $i = 1, 2$ . We prove that  $\text{CMO}^p$  is the dual of  $H^p$ , and in particular that  $\text{BMO}$  is the dual of  $H^1$ , and also that  $H^1$  is the dual of  $\text{VMO}$ .

We develop this theory in the product case with two parameters. The generalization to  $k$  parameters,  $k \in \mathbb{N}$ , is similar to the two-parameter case, while the specialization to one



parameter is immediate. Note the difference from the Littlewood–Paley theory developed in Section 4 above; there it was necessary to develop the one-parameter theory first, then to pass to the product case by iteration.

Fix  $\beta_i \in (0, \eta_i)$  and  $\gamma_i > 0$ , for  $i = 1, 2$ .

For brevity, we denote by  $\mathring{G}$  and  $(\mathring{G})'$  the test function space  $\mathring{G}(\beta_1, \beta_2; \gamma_1, \gamma_2)$  and the space of distributions  $(\mathring{G}(\beta_1, \beta_2; \gamma_1, \gamma_2))'$ , respectively.

We are now ready to introduce the Hardy spaces  $H^p(X_1 \times X_2)$  and the Carleson measure spaces  $\text{CMO}^p(X_1 \times X_2)$ . In this section,  $(x_1, x_2)$  denotes an element of  $X_1 \times X_2$ .

**Definition 5.1.** (Hardy spaces) Suppose  $\max\{\frac{\omega_1}{\omega_1+\eta_1}, \frac{\omega_2}{\omega_2+\eta_2}\} < p \leq 1$ , where  $\omega_i$  is the upper dimension of  $X_i$  for  $i = 1, 2$ . The *Hardy spaces*  $H^p(X_1 \times X_2)$  are defined by

$$H^p(X_1 \times X_2) := \{f \in (\mathring{G})' : \tilde{S}(f) \in L^p(X_1 \times X_2)\},$$

where  $\tilde{S}(f)$  is the discrete product Littlewood–Paley square function as in Definition 4.7.

For  $f \in H^p(X_1 \times X_2)$ , we define  $\|f\|_{H^p(X_1 \times X_2)} := \|\tilde{S}(f)\|_{L^p(X_1 \times X_2)}$ .

For completeness, we note that as in the classical case, for  $1 < p < \infty$  the Hardy space  $H^p(X_1 \times X_2)$  of Definition 5.1 coincides with  $L^p(X_1 \times X_2)$ .

We point out that  $\mathring{G}$  and hence  $H^p(X_1 \times X_2) \cap L^2(X_1 \times X_2)$  are dense in  $H^p(X_1 \times X_2)$ . Indeed, if  $f \in H^p(X_1 \times X_2)$ , then by Theorem 3.4, the functions

$$f_n(x_1, x_2) := \sum_{|k_1|, |k_2| \leq n} \sum_{\alpha_1 \in \mathcal{Y}^{k_1}} \sum_{\alpha_2 \in \mathcal{Y}^{k_2}} \psi_{\alpha_1}^{k_1}(x_1) \psi_{\alpha_2}^{k_2}(x_2) \langle \psi_{\alpha_1}^{k_1} \psi_{\alpha_2}^{k_2}, f \rangle$$

belong to  $\mathring{G}$ . Moreover,

$$\tilde{S}(f - f_n)(x_1, x_2) \leq \left\{ \sum_{|k_1| > n \text{ or } |k_2| > n} \sum_{\alpha_1 \in \mathcal{Y}^{k_1}} \sum_{\alpha_2 \in \mathcal{Y}^{k_2}} \left| \langle \psi_{\alpha_1}^{k_1} \psi_{\alpha_2}^{k_2}, f \rangle \tilde{\chi}_{Q_{\alpha_1}^{k_1}}(x_1) \tilde{\chi}_{Q_{\alpha_2}^{k_2}}(x_2) \right|^2 \right\}^{1/2}.$$

Therefore  $\|\tilde{S}(f - f_n)\|_{L^p(X_1 \times X_2)}$  tends to zero as  $n$  tends to infinity. Hence  $\mathring{G}$  is dense in  $H^p(X_1 \times X_2)$ .

**Definition 5.2.** (Carleson measure spaces, and bounded mean oscillation) Suppose that  $\max\{\frac{\omega_1}{\omega_1+\eta_1}, \frac{\omega_2}{\omega_2+\eta_2}\} < p \leq 1$ , where  $\omega_i$  is the upper dimension of  $X_i$  for  $i = 1, 2$ . We define the *Carleson measure spaces*  $\text{CMO}^p$  in terms of wavelet coefficients by

$$\text{CMO}^p(X_1 \times X_2) := \{f \in (\mathring{G})' : \mathcal{C}_p(f) < L^\infty\},$$

with the quantity  $\mathcal{C}_p(f)$  defined as follows:

$$(5.1) \quad \mathcal{C}_p(f) := \sup_{\Omega} \left\{ \frac{1}{\mu(\Omega)^{\frac{2}{p}-1}} \sum_{\substack{R=Q_{\alpha_1}^{k_1} \times Q_{\alpha_2}^{k_2} \subset \Omega, \\ k_1, k_2 \in \mathbb{Z}, \alpha_1 \in \mathcal{Y}^{k_1}, \alpha_2 \in \mathcal{Y}^{k_2}}} \left| \langle \psi_{\alpha_1}^{k_1} \psi_{\alpha_2}^{k_2}, f \rangle \right|^2 \right\}^{1/2},$$

where  $\Omega$  runs over all open sets in  $X_1 \times X_2$  with finite measure.

The *space BMO of functions of bounded mean oscillation* is defined by

$$\text{BMO}(X_1 \times X_2) := \text{CMO}^1(X_1 \times X_2).$$

The main result in this section is the following.

**Theorem 5.3.** *Suppose  $\max\{\frac{\omega_1}{\omega_1+\eta_1}, \frac{\omega_2}{\omega_2+\eta_2}\} < p \leq 1$ , where  $\omega_i$  is the upper dimension of  $X_i$  for  $i = 1, 2$ . Then the Carleson measure space  $\text{CMO}^p(X_1 \times X_2)$  is the dual of the Hardy space  $H^p(X_1 \times X_2)$ :*

$$(H^p(X_1 \times X_2))' = \text{CMO}^p(X_1 \times X_2).$$

In particular,

$$(H^1(X_1 \times X_2))' = \text{BMO}(X_1 \times X_2).$$

To prove Theorem 5.3, we follow the approach developed in [HLL2]; see also [HLL1]. We first recall the definitions of the product sequence spaces  $s^p$  and  $c^p$  for  $0 < p \leq 1$ . These sequence spaces are discrete analogues of  $H^p(X_1 \times X_2)$  and  $\text{CMO}^p(X_1 \times X_2)$  respectively.

The space  $s^p$  is defined to be the set of sequences  $s = \{s_R\}_R$  of real numbers such that

$$(5.2) \quad \|s\|_{s^p} := \left\| \left\{ \sum_R |\mu(R)^{-1/2} s_R \chi_R(\cdot, \cdot)|^2 \right\}^{1/2} \right\|_{L^p(X_1 \times X_2)} < \infty,$$

where  $R$  runs over all dyadic rectangles in  $X_1 \times X_2$ . The space  $c^p$  is defined to be the set of sequences  $t = \{t_R\}_R$  of real numbers such that

$$(5.3) \quad \|t\|_{c^p} := \sup_{\Omega} \left( \frac{1}{\mu(\Omega)^{\frac{2}{p}-1}} \sum_{R \subset \Omega} |t_R|^2 \right)^{1/2} < \infty,$$

where  $\Omega$  runs over all open sets in  $X_1 \times X_2$  with finite measure, and  $R$  runs over all dyadic rectangles contained in  $\Omega$ .

We emphasize that in the above definitions of  $s^p$  and  $c^p$ , the expression “all dyadic rectangles  $R$ ” indicates the rectangles of the form  $R = Q_{\alpha_1}^{k_1} \times Q_{\alpha_2}^{k_2}$  for all  $k_i \in \mathbb{Z}$  and  $\alpha_i \in \mathcal{X}^{k_i}$  for  $i = 1, 2$ .

The main result about the sequence spaces  $s^p$  and  $c^p$  is the following duality result.

**Proposition 5.4** ([HLL2]). *For  $0 < p \leq 1$ ,  $(s^p)' = c^p$ .*

We now introduce the lifting and projection operators  $T_L$  and  $T_P$ , as follows.

**Definition 5.5.** For  $f \in (\mathring{G})'$ , the *lifting operator*  $T_L$  is defined by

$$(5.4) \quad \{(T_L f)_R\}_R := \{ \langle \psi_{\alpha_1}^{k_1} \psi_{\alpha_2}^{k_2}, f \rangle \}_R,$$

where  $R = Q_{\alpha_1}^{k_1} \times Q_{\alpha_2}^{k_2}$ ,  $k_1, k_2 \in \mathbb{Z}$ ,  $\alpha_1 \in \mathcal{Y}^{k_1}$ ,  $\alpha_2 \in \mathcal{Y}^{k_2}$  are dyadic rectangles in  $X_1 \times X_2$ .

**Definition 5.6.** Given a sequence  $\lambda = \{\lambda_R\}$  of real numbers, we define the associated *projection operator*  $T_P$  by

$$(5.5) \quad T_P(\lambda)(x_1, x_2) := \sum_{\substack{R=Q_{\alpha_1}^{k_1} \times Q_{\alpha_2}^{k_2}, \\ k_1, k_2 \in \mathbb{Z}, \alpha_1 \in \mathcal{Y}^{k_1}, \alpha_2 \in \mathcal{Y}^{k_2}}} \lambda_R \cdot \psi_{\alpha_1}^{k_1}(x_1) \psi_{\alpha_2}^{k_2}(x_2).$$

From the definitions of the lifting and projection operators  $T_L$  and  $T_P$ , it follows that  $f = T_P \circ T_L(f)$  in the sense of the test function space  $\mathring{G}$  and of the distributions  $(\mathring{G})'$ . That is,  $T_P \circ T_L$  is an identity operator on the distributions  $(\mathring{G})'$ .

Next we give two auxiliary results which will be used in establishing the duality in Theorem 5.3.

**Proposition 5.7.** *Suppose  $\max \left\{ \frac{\omega_1}{\omega_1 + \eta_1}, \frac{\omega_2}{\omega_2 + \eta_2} \right\} < p \leq 1$ , where  $\omega_i$  is the upper dimension of  $X_i$  for  $i = 1, 2$ . Then for all  $f \in H^p(X_1 \times X_2)$ , we have*

$$(5.6) \quad \|T_L(f)\|_{s^p} \lesssim \|f\|_{H^p(X_1 \times X_2)}.$$

*In the other direction, for each  $s \in s^p$  we have*

$$(5.7) \quad \|T_P(s)\|_{H^p(X_1 \times X_2)} \lesssim \|s\|_{s^p}.$$

*Proof.* Inequality (5.6) follows directly from the definitions of  $H^p(X_1 \times X_2)$  (Definition 5.1) and the sequence space  $s^p$  (formula (5.2)).

We now prove (5.7). For each  $s \in s^p$ , by the definitions of  $H^p(X_1 \times X_2)$  and  $T_P(s)$ , we have

$$\begin{aligned} \|T_P(s)\|_{H^p(X_1 \times X_2)} &= \|\tilde{S}(T_P(s))\|_{L^p(X_1 \times X_2)} \\ &= \left\| \left\{ \sum_{\substack{R=Q_{\alpha_1}^{k_1} \times Q_{\alpha_2}^{k_2}, \\ k_1, k_2 \in \mathbb{Z}, \alpha_1 \in \mathcal{Y}^{k_1}, \alpha_2 \in \mathcal{Y}^{k_2}}} \left| \left\langle \psi_{\alpha_1}^{k_1} \psi_{\alpha_2}^{k_2}, \sum_{\substack{R'=Q_{\alpha'_1}^{k'_1} \times Q_{\alpha'_2}^{k'_2}, \\ k'_1, k'_2 \in \mathbb{Z}, \alpha'_1 \in \mathcal{Y}^{k'_1}, \alpha'_2 \in \mathcal{Y}^{k'_2}}} s_{R'} \cdot \psi_{\alpha'_1}^{k'_1} \psi_{\alpha'_2}^{k'_2} \right\rangle \right. \right. \\ &\quad \left. \left. \tilde{\chi}_{Q_{\alpha_1}^{k_1}}(x_1) \tilde{\chi}_{Q_{\alpha_2}^{k_2}}(x_2) \right|^2 \right\}^{1/2} \right\|_{L^p(X_1 \times X_2)} \\ &= \left\| \left\{ \sum_{\substack{R=Q_{\alpha_1}^{k_1} \times Q_{\alpha_2}^{k_2}, \\ k_1, k_2 \in \mathbb{Z}, \alpha_1 \in \mathcal{Y}^{k_1}, \alpha_2 \in \mathcal{Y}^{k_2}}} |s_R \cdot \tilde{\chi}_{Q_{\alpha_1}^{k_1}}(x_1) \tilde{\chi}_{Q_{\alpha_2}^{k_2}}(x_2)|^2 \right\}^{1/2} \right\|_{L^p(X_1 \times X_2)} \\ &\leq \|s\|_{s^p}, \end{aligned}$$

where the third equality follows from the orthogonality of the bases  $\{\psi_{\alpha_1}^{k_1}\}$  and  $\{\psi_{\alpha_2}^{k_2}\}$ .  $\square$

**Proposition 5.8.** *Suppose  $\max \left\{ \frac{\omega_1}{\omega_1 + \eta_1}, \frac{\omega_2}{\omega_2 + \eta_2} \right\} < p \leq 1$ , where  $\omega_i$  is the upper dimension of  $X_i$  for  $i = 1, 2$ . For all  $f \in \text{CMO}^p(X_1 \times X_2)$ , we have*

$$(5.8) \quad \|T_L(f)\|_{c^p} \lesssim \mathcal{C}_p(f).$$

*In the other direction, for each  $t \in c^p$ ,*

$$(5.9) \quad \mathcal{C}_p(T_P(t)) \lesssim \|t\|_{c^p}.$$

*Proof.* Inequality (5.8) follows directly from the definitions of  $\text{CMO}^p(X_1 \times X_2)$  (Definition 5.2) and  $c^p$  (formula (5.3)).

We now prove (5.9). For each  $t \in c^p$  we have

$$\begin{aligned} \mathcal{C}_p(T_P(t)) &= \sup_{\Omega} \left\{ \frac{1}{\mu(\Omega)} \sum_{\substack{R=Q_{\alpha_1}^{k_1} \times Q_{\alpha_2}^{k_2} \subset \Omega, \\ k_1, k_2 \in \mathbb{Z}, \alpha_1 \in \mathcal{Y}^{k_1}, \alpha_2 \in \mathcal{Y}^{k_2}}} \left| \langle \psi_{\alpha_1}^{k_1} \psi_{\alpha_2}^{k_2}, T_P(t) \rangle \right|^2 \right\}^{1/2} \\ &= \sup_{\Omega} \left\{ \frac{1}{\mu(\Omega)} \sum_{\substack{R=Q_{\alpha_1}^{k_1} \times Q_{\alpha_2}^{k_2} \subset \Omega, \\ k_1, k_2 \in \mathbb{Z}, \alpha_1 \in \mathcal{Y}^{k_1}, \alpha_2 \in \mathcal{Y}^{k_2}}} |t_R|^2 \right\}^{1/2} \\ &\leq \|t\|_{c^p}, \end{aligned}$$

where the second equality follows from the orthogonality of the bases  $\{\psi_{\alpha_1}^{k_1}\}$  and  $\{\psi_{\alpha_2}^{k_2}\}$ .  $\square$

We would like to point out that thanks to the orthogonality of the wavelet basis from [AH], the proofs given here of (5.7) and (5.9) are much simpler than those given in [HLL2].

We are ready to prove the duality  $(H^p(X_1 \times X_2))' = \text{CMO}^p(X_1 \times X_2)$ .

*Proof of Theorem 5.3.* Suppose  $\max\{\frac{\omega_1}{\omega_1+\eta_1}, \frac{\omega_2}{\omega_2+\eta_2}\} < p \leq 1$ . We first show that there exists a positive constant  $C$  such that for each  $g \in \text{CMO}^p(X_1 \times X_2)$ ,

$$(5.10) \quad |\langle f, g \rangle| \leq C \|f\|_{H^p(X_1 \times X_2)} \mathcal{C}_p(g)$$

for all  $f \in \mathring{G}$ . It follows that  $\text{CMO}^p(X_1 \times X_2) \subset (H^p(X_1 \times X_2))'$ , since  $\mathring{G}$  is dense in  $H^p(X_1 \times X_2)$ .

To prove inequality (5.10), for each  $f \in \mathring{G}$  and  $g \in \text{CMO}^p(X_1 \times X_2)$ , by the reproducing formula (3.26) we have

$$\begin{aligned} \langle f, g \rangle &= \sum_{k_1, k_2 \in \mathbb{Z}, \alpha_1 \in \mathcal{Y}^{k_1}, \alpha_2 \in \mathcal{Y}^{k_2}} \langle \psi_{\alpha_1}^{k_1} \psi_{\alpha_2}^{k_2}, f \rangle \langle \psi_{\alpha_1}^{k_1} \psi_{\alpha_2}^{k_2}, g \rangle \\ &= \sum_R (T_L(f))_R \cdot (T_L(g))_R. \end{aligned}$$

where  $T_L(f)$  and  $T_L(g)$  are the lifting operators as in Definition 5.5.

Then, by Propositions 5.7 and 5.8, we obtain

$$|\langle f, g \rangle| \leq |\langle T_L(f), T_L(g) \rangle| \leq C \|f\|_{H^p(X_1 \times X_2)} \mathcal{C}_p(g).$$

Conversely, suppose  $l \in (H^p(X_1 \times X_2))'$ . Let  $l_1 := l \circ T_P$ . By Proposition 5.7, we see that  $l_1 \in (s^p)'$ , since for each  $s \in s^p$ ,  $|l_1(s)| = |l(T_P(s))| \leq C \|l\| \|T_P(s)\|_{H^p(X_1 \times X_2)} \leq C \|l\| \|s\|_{s^p}$ . Now we have

$$l(g) = l \circ T_P \circ T_L(g) = l_1(T_L(g))$$

for each  $g \in \mathring{G}$ . So by Proposition 5.4, there exists  $t \in c^p$  such that  $l_1(s) = \langle t, s \rangle$  for all  $s \in s^p$  and  $\|t\|_{c^p} \sim \|l_1\| \lesssim \|l\|$ . Hence

$$l(g) = \langle t, T_L(g) \rangle = \langle T_P(t), g \rangle.$$

By Definition 5.2 and Proposition 5.8, we obtain that  $\|T_P(t)\|_{\text{CMO}^p(X_1 \times X_2)} \lesssim \|t\|_{c^p} \lesssim \|l\|$ . Hence  $(H^p(X_1 \times X_2))' \subset \text{CMO}^p(X_1 \times X_2)$ .  $\square$

Now we introduce the space of functions of vanishing mean oscillation.

**Definition 5.9.** (Vanishing mean oscillation) We define the *space*  $\text{VMO}(X_1 \times X_2)$  of functions of vanishing mean oscillation to be the subspace of  $\text{BMO}(X_1 \times X_2)$  consisting of those  $f \in \text{BMO}(X_1 \times X_2)$  satisfying the three properties

- (a)  $\lim_{\delta \rightarrow 0} \sup_{\Omega: \mu(\Omega) < \delta} \left\{ \frac{1}{\mu(\Omega)} \sum_{\substack{R=Q_{\alpha_1}^{k_1} \times Q_{\alpha_2}^{k_2} \subset \Omega \\ k_1, k_2 \in \mathbb{Z}, \alpha_1 \in \mathcal{Y}^{k_1}, \alpha_2 \in \mathcal{Y}^{k_2}}} |\langle \psi_{\alpha_1}^{k_1} \psi_{\alpha_2}^{k_2}, f \rangle|^2 \right\}^{1/2} = 0;$
- (b)  $\lim_{N \rightarrow \infty} \sup_{\Omega: \text{diam}(\Omega) > N} \left\{ \frac{1}{\mu(\Omega)} \sum_{\substack{R=Q_{\alpha_1}^{k_1} \times Q_{\alpha_2}^{k_2} \subset \Omega, \\ k_1, k_2 \in \mathbb{Z}, \alpha_1 \in \mathcal{Y}^{k_1}, \alpha_2 \in \mathcal{Y}^{k_2}}} |\langle \psi_{\alpha_1}^{k_1} \psi_{\alpha_2}^{k_2}, f \rangle|^2 \right\}^{1/2} = 0;$
- (c)  $\lim_{N \rightarrow \infty} \sup_{\Omega: \Omega \subset (B(x_1, N) \times B(x_2, N))^c} \left\{ \frac{1}{\mu(\Omega)} \sum_{\substack{R=Q_{\alpha_1}^{k_1} \times Q_{\alpha_2}^{k_2} \subset \Omega \\ k_1, k_2 \in \mathbb{Z}, \alpha_1 \in \mathcal{Y}^{k_1}, \alpha_2 \in \mathcal{Y}^{k_2}}} |\langle \psi_{\alpha_1}^{k_1} \psi_{\alpha_2}^{k_2}, f \rangle|^2 \right\}^{1/2} = 0;$

$$\left\{ \frac{1}{\mu(\Omega)} \sum_{\substack{R=Q_{\alpha_1}^{k_1} \times Q_{\alpha_2}^{k_2} \subset \Omega \\ k_1, k_2 \in \mathbb{Z}, \alpha_1 \in \mathcal{D}^{k_1}, \alpha_2 \in \mathcal{D}^{k_2}}} |\langle \psi_{\alpha_1}^{k_1} \psi_{\alpha_2}^{k_2}, f \rangle|^2 \right\}^{1/2} = 0, \text{ where}$$

$x_1$  and  $x_2$  are arbitrary fixed points in  $X_1$  and  $X_2$ , respectively.

We now show the duality of  $\text{VMO}(X_1 \times X_2)$  with  $H^1(X_1 \times X_2)$ .

**Theorem 5.10.** *The Hardy space  $H^1(X_1 \times X_2)$  is the dual of  $\text{VMO}(X_1 \times X_2)$ :*

$$(\text{VMO}(X_1 \times X_2))' = H^1(X_1 \times X_2).$$

*Proof.* The proof of this theorem is similar to the proof of the duality between  $\text{VMO}$  and  $H^1$  on Euclidean space given in Section 5 of [LTW]. Following [LTW], we only sketch the main steps of the proof. First, we use  $FW$  to denote the set of finite linear combinations of terms of the form  $\{\psi_{\alpha_1}^{k_1} \cdot \psi_{\alpha_2}^{k_2}\}$ , where  $\{\psi_{\alpha_i}^{k_i}\}$  are wavelets on  $X_i$ ,  $i = 1, 2$ , as in Theorem 2.2. Second, from Definition 5.9, we obtain that  $\text{VMO}(X_1 \times X_2)$  is the closure of  $FW$  in the  $\text{BMO}(X_1 \times X_2)$  norm.

The inclusion  $H^1(X_1 \times X_2) \subset (\text{VMO}(X_1 \times X_2))'$  follows from the duality of  $H^1(X_1 \times X_2)$  with  $\text{BMO}(X_1 \times X_2)$ , which was shown in Theorem 5.3. The reverse containment follows from the fact that  $FW$  is dense in  $H^1(X_1 \times X_2)$  in terms of the  $H^1(X_1 \times X_2)$  norm and from the following inequality: for  $f \in FW$ ,

$$\|f\|_{H^1(X_1 \times X_2)} \leq C \sup_{\substack{b \in \text{VMO}(X_1 \times X_2), \\ \|b\|_{\text{BMO}(X_1 \times X_2)} = 1}} |\langle b, f \rangle|. \quad \square$$

## 6. CALDERÓN–ZYGmund DECOMPOSITION AND INTERPOLATION ON HARDY SPACES

In this section we provide the Calderón–Zygmund decomposition and prove an interpolation theorem on  $H^p(X_1 \times X_2)$ . Note that  $H^p(X_1 \times X_2) = L^p(X_1 \times X_2)$  for  $1 < p < \infty$ . In this section,  $(x_1, x_2)$  denotes an element of  $X_1 \times X_2$ .

**Theorem 6.1.** *Let  $\max\{\frac{\omega_1}{\omega_1 + \eta_1}, \frac{\omega_2}{\omega_2 + \eta_2}\} < p_2 \leq 1$ , where  $\omega_i$  is the upper dimension of  $X_i$  for  $i = 1, 2$ . Suppose  $p_2 < p < p_1 < \infty$ ,  $\alpha > 0$ , and  $f \in H^p(X_1 \times X_2)$ . Then we may write*

$$f(x_1, x_2) = g(x_1, x_2) + b(x_1, x_2),$$

where

$$g \in H^{p_1}(X_1 \times X_2) \quad \text{and} \quad b \in H^{p_2}(X_1 \times X_2)$$

are such that  $\|g\|_{H^{p_1}(X_1 \times X_2)}^{p_1} \leq C \alpha^{p_1 - p} \|f\|_{H^p(X_1 \times X_2)}^p$  and  $\|b\|_{H^{p_2}(X_1 \times X_2)}^{p_2} \leq C \alpha^{p_2 - p} \|f\|_{H^p(X_1 \times X_2)}^p$ . Here  $C$  is an absolute constant.

**Theorem 6.2.** *Suppose  $\max\{\frac{\omega_1}{\omega_1 + \eta_1}, \frac{\omega_2}{\omega_2 + \eta_2}\} < p_2 < p_1 < \infty$ , where  $\omega_i$  is the upper dimension of  $X_i$  for  $i = 1, 2$ . Then the following two assertions hold.*

- (a) *Let  $T$  be a linear operator that is bounded from  $H^{p_2}(X_1 \times X_2)$  to  $L^{p_2}(X_1 \times X_2)$  and from  $H^{p_1}(X_1 \times X_2)$  to  $L^{p_1}(X_1 \times X_2)$ . Then  $T$  is bounded from  $H^p(X_1 \times X_2)$  to  $L^p(X_1 \times X_2)$  for all  $p$  with  $p_2 < p < p_1$ .*
- (b) *Suppose  $T$  is bounded on  $H^{p_2}(X_1 \times X_2)$  and on  $H^{p_1}(X_1 \times X_2)$ . Then  $T$  is bounded on  $H^p(X_1 \times X_2)$  for all  $p$  with  $p_2 < p < p_1$ .*

We first prove Theorem 6.1.

**Proof of Theorem 6.1.** Suppose that  $f \in H^p(X_1 \times X_2)$  and  $\alpha > 0$ . Let  $\Omega_\ell := \{(x_1, x_2) \in X_1 \times X_2 : \tilde{S}(f)(x_1, x_2) > \alpha 2^\ell\}$ , where  $\tilde{S}(f)$  is the discrete product square function defined in (4.10).

Let

$$\mathcal{R}_0 := \left\{ R = Q_{\alpha_1}^{k_1} \times Q_{\alpha_2}^{k_2} : k_1, k_2 \in \mathbb{Z}, \alpha_1 \in \mathcal{Y}^{k_1}, \alpha_2 \in \mathcal{Y}^{k_2}, \mu(R \cap \Omega_0) < \frac{1}{2A_0} \mu(R) \right\}$$

and for  $\ell \geq 1$

$$\mathcal{R}_\ell := \left\{ R = Q_{\alpha_1}^{k_1} \times Q_{\alpha_2}^{k_2}, k_1, k_2 \in \mathbb{Z}, \alpha_1 \in \mathcal{Y}^{k_1}, \alpha_2 \in \mathcal{Y}^{k_2} \right. \\ \left. \text{such that } \mu(R \cap \Omega_{\ell-1}) \geq \frac{1}{2A_0} \mu(R) \text{ and } \mu(R \cap \Omega_\ell) < \frac{1}{2A_0} \mu(R) \right\}.$$

Applying the wavelet reproducing formula from Theorem 3.11, we have

$$\begin{aligned} f(x_1, x_2) &= \sum_{k_1} \sum_{\alpha_1 \in \mathcal{Y}^{k_1}} \sum_{k_2} \sum_{\alpha_2 \in \mathcal{Y}^{k_2}} \langle f, \psi_{\alpha_1}^{k_1} \psi_{\alpha_2}^{k_2} \rangle \psi_{\alpha_1}^{k_1}(x_1) \psi_{\alpha_2}^{k_2}(x_2) \\ &= \sum_{\ell \geq 1} \sum_{R=Q_{\alpha_1}^{k_1} \times Q_{\alpha_2}^{k_2} \in \mathcal{R}_\ell} \langle f, \psi_{\alpha_1}^{k_1} \psi_{\alpha_2}^{k_2} \rangle \psi_{\alpha_1}^{k_1}(x_1) \psi_{\alpha_2}^{k_2}(x_2) \\ &\quad + \sum_{R=Q_{\alpha_1}^{k_1} \times Q_{\alpha_2}^{k_2} \in \mathcal{R}_0} \langle f, \psi_{\alpha_1}^{k_1} \psi_{\alpha_2}^{k_2} \rangle \psi_{\alpha_1}^{k_1}(x_1) \psi_{\alpha_2}^{k_2}(x_2) \\ &=: b(x, y) + g(x, y). \end{aligned}$$

When  $p_1 > 1$ , the  $L^p(X_1 \times X_2)$ ,  $1 < p < \infty$ , estimate for the Littlewood–Paley square function implies that

$$\|g\|_{L^{p_1}(X_1 \times X_2)} \leq C \left\| \left\{ \sum_{R=Q_{\alpha_1}^{k_1} \times Q_{\alpha_2}^{k_2} \in \mathcal{R}_0} \left| \langle \psi_{\alpha_1}^{k_1} \psi_{\alpha_2}^{k_2}, f \rangle \tilde{\chi}_{Q_{\alpha_1}^{k_1}}(x_1) \tilde{\chi}_{Q_{\alpha_2}^{k_2}}(x_2) \right|^2 \right\}^{1/2} \right\|_{L^{p_1}(X_1 \times X_2)}.$$

Next, we estimate  $\|g\|_{H^{p_1}(X_1 \times X_2)}$  when  $\max\{\frac{\omega_1}{\omega_1 + \eta_1}, \frac{\omega_2}{\omega_2 + \eta_2}\} < p_1 \leq 1$ . We estimate the  $H^{p_1}(X_1 \times X_2)$  norm directly. To this end, using the wavelet coefficients of  $g$ , we observe that

$$\begin{aligned} \|g\|_{H^{p_1}(X_1 \times X_2)} &\leq \left\| \left\{ \sum_{k'_1} \sum_{\alpha'_1 \in \mathcal{Y}^{k'_1}} \sum_{k'_2} \sum_{\alpha'_2 \in \mathcal{Y}^{k'_2}} \left| \langle \psi_{\alpha'_1}^{k'_1} \psi_{\alpha'_2}^{k'_2}, g \rangle \tilde{\chi}_{Q_{\alpha'_1}^{k'_1}}(x_1) \tilde{\chi}_{Q_{\alpha'_2}^{k'_2}}(x_2) \right|^2 \right\}^{1/2} \right\|_{L^{p_1}(X_1 \times X_2)} \\ &\leq C \left\| \left\{ \sum_{R=Q_{\alpha_1}^{k_1} \times Q_{\alpha_2}^{k_2} \in \mathcal{R}_0} \left| \langle \psi_{\alpha_1}^{k_1} \psi_{\alpha_2}^{k_2}, f \rangle \tilde{\chi}_{Q_{\alpha_1}^{k_1}}(x_1) \tilde{\chi}_{Q_{\alpha_2}^{k_2}}(x_2) \right|^2 \right\}^{1/2} \right\|_{L^{p_1}(X_1 \times X_2)}. \end{aligned}$$

Thus for all  $p_1$  with  $\max\{\frac{\omega_1}{\omega_1 + \eta_1}, \frac{\omega_2}{\omega_2 + \eta_2}\} < p_1 < \infty$ , we have

$$\|g\|_{H^{p_1}(X_1 \times X_2)} \leq C \left\| \left\{ \sum_{R=Q_{\alpha_1}^{k_1} \times Q_{\alpha_2}^{k_2} \in \mathcal{R}_0} \left| \langle \psi_{\alpha_1}^{k_1} \psi_{\alpha_2}^{k_2}, f \rangle \tilde{\chi}_{Q_{\alpha_1}^{k_1}}(x_1) \tilde{\chi}_{Q_{\alpha_2}^{k_2}}(x_2) \right|^2 \right\}^{1/2} \right\|_{L^{p_1}(X_1 \times X_2)}.$$

**Claim 1:** We claim that

$$\begin{aligned} &\int_{\tilde{S}(f)(x_1, x_2) \leq \alpha} \tilde{S}(f)(x_1, x_2)^{p_1} d\mu_1(x_1) d\mu_2(x_2) \\ &\geq C \left\| \left\{ \sum_{R=Q_{\alpha_1}^{k_1} \times Q_{\alpha_2}^{k_2} \in \mathcal{R}_0} \left| \langle \psi_{\alpha_1}^{k_1} \psi_{\alpha_2}^{k_2}, f \rangle \tilde{\chi}_{Q_{\alpha_1}^{k_1}}(x_1) \tilde{\chi}_{Q_{\alpha_2}^{k_2}}(x_2) \right|^2 \right\}^{1/2} \right\|_{L^{p_1}(X_1 \times X_2)}. \end{aligned}$$

This implies that

$$\begin{aligned}
\|g\|_{H^{p_1}(X_1 \times X_2)} &\leq C \int_{\tilde{S}(f)(x_1, x_2) \leq \alpha} \tilde{S}(f)(x_1, x_2)^{p_1} d\mu_1(x_1) d\mu_2(x_2) \\
&\leq C \alpha^{p_1-p} \int_{\tilde{S}(f)(x_1, x_2) \leq \alpha} \tilde{S}(f)(x_1, x_2)^p d\mu_1(x_1) d\mu_2(x_2) \\
&\leq C \alpha^{p_1-p} \|f\|_{H^p(X_1 \times X_2)}^p.
\end{aligned}$$

To show Claim 1, we choose  $0 < q < p_1$  and  $q < 2$ , and observe that

$$\begin{aligned}
&\int_{\tilde{S}(f)(x_1, x_2) \leq \alpha} \tilde{S}(f)(x_1, x_2)^{p_1} d\mu_1(x_1) d\mu_2(x_2) \\
&= \int_{\Omega_0^c} \left\{ \sum_{k_1} \sum_{\alpha_1 \in \mathcal{Y}^{k_1}} \sum_{k_2} \sum_{\alpha_2 \in \mathcal{Y}^{k_2}} \left| \langle \psi_{\alpha_1}^{k_1} \psi_{\alpha_2}^{k_2}, f \rangle \tilde{\chi}_{Q_{\alpha_1}^{k_1}}(x_1) \tilde{\chi}_{Q_{\alpha_2}^{k_2}}(x_2) \right|^2 \right\}^{p_1/2} d\mu_1(x_1) d\mu_2(x_2) \\
&\geq C \int_{\Omega_0^c} \left\{ \sum_{R=Q_{\alpha_1}^{k_1} \times Q_{\alpha_2}^{k_2} \in \mathcal{R}_0} \left| \langle \psi_{\alpha_1}^{k_1} \psi_{\alpha_2}^{k_2}, f \rangle \tilde{\chi}_{Q_{\alpha_1}^{k_1}}(x_1) \tilde{\chi}_{Q_{\alpha_2}^{k_2}}(x_2) \right|^2 \right\}^{\frac{p_1}{2}} d\mu_1(x_1) d\mu_2(x_2) \\
&= C \int_{X_1 \times X_2} \left\{ \sum_{R=Q_{\alpha_1}^{k_1} \times Q_{\alpha_2}^{k_2} \in \mathcal{R}_0} \left| \langle \psi_{\alpha_1}^{k_1} \psi_{\alpha_2}^{k_2}, f \rangle \tilde{\chi}_{Q_{\alpha_1}^{k_1}}(x_1) \tilde{\chi}_{Q_{\alpha_2}^{k_2}}(x_2) \chi_{\Omega_0^c}(x_1, x_2) \right|^2 \right\}^{\frac{p_1}{2}} d\mu_1(x_1) d\mu_2(x_2) \\
&\geq C \int_{X_1 \times X_2} \left[ \left\{ \sum_{R=Q_{\alpha_1}^{k_1} \times Q_{\alpha_2}^{k_2} \in \mathcal{R}_0} \left( M_s(\langle \psi_{\alpha_1}^{k_1} \psi_{\alpha_2}^{k_2}, f \rangle^q \mu(R)^{-q} \chi_{R \cap \Omega_0^c})(x_1, x_2) \right)^{\frac{2}{q}} \right\}^{\frac{q}{2}} \right]^{\frac{p_1}{q}} d\mu_1(x_1) d\mu_2(x_2) \\
&\geq C \int_{X_1 \times X_2} \left\{ \sum_{R=Q_{\alpha_1}^{k_1} \times Q_{\alpha_2}^{k_2} \in \mathcal{R}_0} \left| \langle \psi_{\alpha_1}^{k_1} \psi_{\alpha_2}^{k_2}, f \rangle \tilde{\chi}_{Q_{\alpha_1}^{k_1}}(x_1) \tilde{\chi}_{Q_{\alpha_2}^{k_2}}(x_2) \right|^2 \right\}^{\frac{p_1}{2}} d\mu_1(x_1) d\mu_2(x_2),
\end{aligned}$$

where in the last inequality we have used the fact that  $\mu(\Omega_0^c \cap R) \geq \frac{1}{2} \mu(R)$  for  $R \in \mathcal{R}_0$ , and thus  $\chi_R(x_1, x_2) \leq 2^{\frac{1}{q}} M_s(\chi_{R \cap \Omega_0^c})^{\frac{1}{q}}(x_1, x_2)$ , and in the second to last inequality we have used the vector-valued Fefferman–Stein inequality for strong maximal functions:

$$\left\| \left\{ \sum_{k=1}^{\infty} M_s(f_k)^r \right\}^{\frac{1}{r}} \right\|_{L^p(X_1 \times X_2)} \leq C \left\| \left\{ \sum_{k=1}^{\infty} |f_k|^r \right\}^{\frac{1}{r}} \right\|_{L^p(X_1 \times X_2)},$$

with the exponents  $r = 2/q > 1$  and  $p = p_1/q > 1$ . Thus the claim follows.

Let  $\widetilde{\Omega}_\ell$  be the enlargement of the set  $\Omega_\ell$  given by  $\widetilde{\Omega}_\ell := \{(x_1, x_2) \in X_1 \times X_2 : M_s(\chi_{\Omega_\ell}) > (2A_0)^{-1}\}$ .

**Claim 2:** For  $p_2 \leq 1$ ,

$$\left\| \sum_{R=Q_{\alpha_1}^{k_1} \times Q_{\alpha_2}^{k_2} \in \mathcal{R}_\ell} \langle f, \psi_{\alpha_1}^{k_1} \psi_{\alpha_2}^{k_2} \rangle \psi_{\alpha_1}^{k_1}(x_1) \psi_{\alpha_2}^{k_2}(x_2) \right\|_{H^{p_2}(X_1 \times X_2)}^{p_2} \leq C (2^\ell \alpha)^{p_2} \mu(\widetilde{\Omega}_{\ell-1}).$$

Claim 2 implies that

$$\begin{aligned}
\|b\|_{H^{p_2}(X_1 \times X_2)}^{p_2} &\leq \sum_{\ell \geq 1} (2^\ell \alpha)^{p_2} \mu(\widetilde{\Omega}_{\ell-1}) \leq C \sum_{\ell \geq 1} (2^\ell \alpha)^{p_2} \mu(\Omega_{\ell-1}) \\
&\leq C \int_{\tilde{S}(f)(x_1, x_2) > \alpha} \tilde{S}(f)(x_1, x_2)^{p_2} d\mu_1(x_1) d\mu_2(x_2)
\end{aligned}$$

$$\begin{aligned}
&\leq C\alpha^{p_2-p} \int_{\tilde{S}(f)(x_1, x_2) > \alpha} \tilde{S}(f)(x_1, x_2)^p d\mu_1(x_1) d\mu_2(x_2) \\
&\leq C\alpha^{p_2-p} \|f\|_{H^p(X_1 \times X_2)}^p.
\end{aligned}$$

To show Claim 2, note that

$$\begin{aligned}
&\left\| \sum_{R=Q_{\alpha_1}^{k_1} \times Q_{\alpha_2}^{k_2} \in \mathcal{R}_\ell} \langle f, \psi_{\alpha_1}^{k_1} \psi_{\alpha_2}^{k_2} \rangle \psi_{\alpha_1}^{k_1}(x_1) \psi_{\alpha_2}^{k_2}(x_2) \right\|_{H^{p_2}(X_1 \times X_2)}^{p_2} \\
&\leq C \left\| \left\{ \sum_{R=Q_{\alpha_1}^{k_1} \times Q_{\alpha_2}^{k_2} \in \mathcal{R}_\ell} \left| \langle \psi_{\alpha_1}^{k_1} \psi_{\alpha_2}^{k_2}, f \rangle \tilde{\chi}_{Q_{\alpha_1}^{k_1}}(x_1) \tilde{\chi}_{Q_{\alpha_2}^{k_2}}(x_2) \right|^2 \right\}^{1/2} \right\|_{L^{p_2}(X_1 \times X_2)}.
\end{aligned}$$

Then

$$\begin{aligned}
&\sum_{\ell=1}^{\infty} (2^\ell \alpha)^{p_2} \mu(\tilde{\Omega}_{\ell-1}) \\
&\geq \int_{\tilde{\Omega}_{\ell-1} \setminus \Omega_\ell} \tilde{S}(f)^{p_2}(x_1, x_2) d\mu_1(x_1) d\mu_2(x_2) \\
&= \int_{\tilde{\Omega}_{\ell-1} \setminus \Omega_\ell} \left\{ \sum_{k_1} \sum_{\alpha_1 \in \mathcal{Y}^{k_1}} \sum_{k_2} \sum_{\alpha_2 \in \mathcal{Y}^{k_2}} \left| \langle \psi_{\alpha_1}^{k_1} \psi_{\alpha_2}^{k_2}, f \rangle \tilde{\chi}_{Q_{\alpha_1}^{k_1}}(x_1) \tilde{\chi}_{Q_{\alpha_2}^{k_2}}(x_2) \right|^2 \right\}^{\frac{p_2}{2}} d\mu_1(x_1) d\mu_2(x_2) \\
&\geq \int_{X_1 \times X_2} \left\{ \sum_{R=Q_{\alpha_1}^{k_1} \times Q_{\alpha_2}^{k_2} \in \mathcal{R}_0} \left| \langle \psi_{\alpha_1}^{k_1} \psi_{\alpha_2}^{k_2}, f \rangle \tilde{\chi}_{Q_{\alpha_1}^{k_1}}(x_1) \tilde{\chi}_{Q_{\alpha_2}^{k_2}}(x_2) \chi_{\tilde{\Omega}_{\ell-1} \setminus \Omega_\ell}(x_1, x_2) \right|^2 \right\}^{\frac{p_2}{2}} d\mu_1(x_1) d\mu_2(x_2) \\
&\geq C \int_{X_1 \times X_2} \left\{ \sum_{R=Q_{\alpha_1}^{k_1} \times Q_{\alpha_2}^{k_2} \in \mathcal{R}_0} \left| \langle \psi_{\alpha_1}^{k_1} \psi_{\alpha_2}^{k_2}, f \rangle \tilde{\chi}_{Q_{\alpha_1}^{k_1}}(x_1) \tilde{\chi}_{Q_{\alpha_2}^{k_2}}(x_2) \right|^2 \right\}^{\frac{p_2}{2}} d\mu_1(x_1) d\mu_2(x_2),
\end{aligned}$$

where the last inequality follows from the fact that if  $R \in \mathcal{R}_\ell$  then  $R \subset \tilde{\Omega}_{\ell-1}$ , and therefore  $\mu(R \cap (\tilde{\Omega}_{\ell-1} \setminus \Omega_\ell)) > \frac{1}{2}\mu(R)$ . This establishes Claim 2. Hence, the proof of Theorem 6.1 is complete.  $\square$

We end our paper by proving the interpolation theorem on Hardy spaces  $H^p(X_1 \times X_2)$ .

*Proof of Theorem 6.2.* (a) Suppose that  $T$  is bounded from  $H^{p_2}(X_1 \times X_2)$  to  $L^{p_2}(X_1 \times X_2)$  and from  $H^{p_1}(X_1 \times X_2)$  to  $L^{p_1}(X_1 \times X_2)$ . For each given  $\lambda > 0$  and  $f \in H^p(X_1 \times X_2)$ , by the Calderón–Zygmund decomposition we may write

$$f(x_1, x_2) = g(x_1, x_2) + b(x_1, x_2)$$

with

$$\|g\|_{H^{p_1}(X_1 \times X_2)}^{p_1} \leq C\lambda^{p_1-p} \|f\|_{H^p(X_1 \times X_2)}^p \quad \text{and} \quad \|b\|_{H^{p_2}(X_1 \times X_2)}^{p_2} \leq C\lambda^{p_2-p} \|f\|_{H^p(X_1 \times X_2)}^p.$$

Moreover, we have proved the estimates

$$\|g\|_{H^{p_1}(X_1 \times X_2)}^{p_1} \leq C \int_{\tilde{S}(f)(x_1, x_2) \leq \alpha} \tilde{S}(f)^{p_1}(x_1, x_2) d\mu_1(x_1) d\mu_2(x_2)$$

and

$$\|b\|_{H^{p_2}(X_1 \times X_2)}^{p_2} \leq C \int_{\tilde{S}(f)(x_1, x_2) > \alpha} \tilde{S}(f)^{p_2}(x_1, x_2) d\mu_1(x_1) d\mu_2(x_2),$$



which imply that

$$\begin{aligned}
\|Tf\|_{L^p(X_1 \times X_2)}^p &= p \int_0^\infty \alpha^{p-1} \mu(\{(x_1, x_2) : |Tf(x_1, x_2)| > \alpha\}) d\alpha \\
&\leq p \int_0^\infty \alpha^{p-1} \mu(\{(x_1, x_2) : |Tg(x_1, x_2)| > \alpha/2\}) d\alpha \\
&\quad + p \int_0^\infty \alpha^{p-1} \mu(\{(x_1, x_2) : |Tb(x_1, x_2)| > \alpha/2\}) d\alpha \\
&\leq p \int_0^\infty \alpha^{p-p_1-1} \int_{\tilde{S}(f)(x_1, x_2) \leq \alpha} \tilde{S}(f)^{p_1}(x_1, x_2) d\mu_1(x_1) d\mu_2(x_2) d\alpha \\
&\quad + p \int_0^\infty \alpha^{p-p_2-1} \int_{\tilde{S}(f)(x_1, x_2) > \alpha} \tilde{S}(f)^{p_2}(x_1, x_2) d\mu_1(x_1) d\mu_2(x_2) d\alpha \\
&\leq C \|f\|_{H^p(X_1 \times X_2)}^p
\end{aligned}$$

for all  $p$  with  $p_2 < p < p_1$ . Hence  $T$  is bounded from  $H^p(X_1 \times X_2)$  to  $L^p(X_1 \times X_2)$ , as required.

(b) We turn to the second assertion. For each given  $\lambda > 0$  and  $f \in H^p(X_1 \times X_2)$ , by the Calderón–Zygmund decomposition again we have

$$\begin{aligned}
&\mu(\{(x_1, x_2) : |\tilde{S}(Tf)(x_1, x_2)| > \alpha\}) \\
&\leq \mu(\{(x_1, x_2) : |\tilde{S}(Tg)(x_1, x_2)| > \frac{\alpha}{2}\}) + \mu(\{(x_1, x_2) : |\tilde{S}(Tb)(x_1, x_2)| > \frac{\alpha}{2}\}) \\
&\leq C\alpha^{-p_1} \|Tg\|_{H^{p_1}(X_1 \times X_2)}^{p_1} + C\alpha^{-p_2} \|Tb\|_{H^{p_2}(X_1 \times X_2)}^{p_2} \\
&\leq C\alpha^{-p_1} \|g\|_{H^{p_1}(X_1 \times X_2)}^{p_1} + C\alpha^{-p_2} \|b\|_{H^{p_2}(X_1 \times X_2)}^{p_2} \\
&\leq C\alpha^{-p_1} \int_{\tilde{S}(f)(x_1, x_2) \leq \alpha} \tilde{S}(f)^{p_1}(x_1, x_2) d\mu_1(x_1) d\mu_2(x_2) \\
&\quad + C\alpha^{-p_2} \int_{\tilde{S}(f)(x_1, x_2) > \alpha} \tilde{S}(f)^{p_2}(x_1, x_2) d\mu_1(x_1) d\mu_2(x_2).
\end{aligned}$$

Therefore  $\|\tilde{S}(Tf)\|_{L^p(X_1 \times X_2)} \leq C \|\tilde{S}(f)\|_{H^p(X_1 \times X_2)}$ . Hence  $\|Tf\|_{H^p(X_1 \times X_2)} \leq C \|f\|_{H^p(X_1 \times X_2)}$  for all  $p$  with  $p_2 < p < p_1$ , as required.  $\square$

## REFERENCES

- [AH] P. Auscher and T. Hytönen, *Orthonormal bases of regular wavelets in spaces of homogeneous type*, Appl. Comput. Harmon. Anal. **34** (2013), 266–296.
- [AH2] P. Auscher and T. Hytönen, *Addendum to: Orthonormal bases of regular wavelets in spaces of homogeneous type*, <http://arxiv.org/pdf/1503.05397.pdf>, 2015.
- [C] A. Calderón, *Intermediate spaces and interpolation, the complex method*, Studia Math. **24** (1964), 113–190.
- [CF1] S.-Y. A. Chang and R. Fefferman, *A continuous version of the duality of  $H^1$  and BMO on the bi-disc*, Ann. Math. **112** (1980), 179–201.
- [CF2] S.-Y. A. Chang and R. Fefferman, *The Calderón–Zygmund decomposition on product domains*, Amer. J. Math. **104** (1982), 445–468.
- [Chr] M. Christ, *A  $T(b)$  theorem with remarks on analytic capacity and the Cauchy integral*, Colloq. Math. **60/61** (1990), no. 2, 601–628.
- [CW1] R.R. Coifman and G. Weiss, *Analyse harmonique non-commutative sur certains espaces homogènes. Étude de certaines intégrales singulières*, Lecture Notes in Math. **242**, Springer-Verlag, Berlin, 1971.
- [CW2] R.R. Coifman and G. Weiss, *Extensions of Hardy spaces and their use in analysis*, Bull. Amer. Math. Soc. **83** (1977), 569–645.

- [DJS] G. David, J.-L. Journé and S. Semmes, *Calderón–Zygmund operators, para-accretive functions and interpolation*, Rev. Mat. Iberoamericana, **1** (1985), no. 4, 1–56.
- [DH] D. Deng and Y. Han, *Harmonic analysis on spaces of homogeneous type*, Lecture Notes in Math., vol. 1966, Springer-Verlag, Berlin, 2009, with a preface by Yves Meyer.
- [FS] C. Fefferman and E.M. Stein,  *$H^p$  spaces of several variables*, Acta Math. **129** (1972), 137–195.
- [FJ] M. Frazier and B. Jawerth, *A discrete transform and decomposition of distribution spaces*, J. Funct. Anal. **93** (1990), 34–170.
- [GS] R. Gundy and E.M. Stein,  *$H^p$  theory for the poly-disc*, Proc. Nat. Acad. Sci. **76** (1972), 1026–1029.
- [H1] Y. Han, *Calderón-type reproducing formula and the Tb theorem*, Rev. Mat. Iberoamericana **10** (1994), 51–91.
- [H2] Y. Han, *Plancherel–Pôlya type inequality on spaces of homogeneous type and its applications*, Proc. Amer. Math. Soc. **126** (1998), no. 11, 3315–3327.
- [HLL1] Y. Han, J. Li and G. Lu, *Duality of multiparameter Hardy spaces  $H^p$  on spaces of homogeneous type*, Ann. Scuola Norm. Super. Pisa Cl. Sci. (5) Vol. **9** (2010), no. 4, 645–685.
- [HLL2] Y. Han, J. Li and G. Lu, *Multiparameter Hardy space theory on Carnot–Carathéodory spaces and product spaces of homogeneous type*, Trans. Amer. Math. Soc. **365** (2013), no. 1, 319–360.
- [HMY] Y. Han, D. Müller and D. Yang, *A theory of Besov and Triebel–Lizorkin spaces on metric measure spaces modeled on Carnot–Carathéodory spaces*, Abstr. Appl. Anal., Vol. 2008, Article ID 893409. 250 pages.
- [HS] Y. Han and E.T. Sawyer, *Littlewood–Paley theory on spaces of homogeneous type and the classical function spaces*, Mem. Amer. Math. Soc. **110** (1994), no. 530, vi + 126 pp.
- [HK] T. Hytönen and A. Kairema, *Systems of dyadic cubes in a doubling metric space*, Colloq. Math. **126** (2012), no. 1, 1–33.
- [LTW] M. Lacey, E. Terwilleger and B. Wick, *Remarks on product VMO*, Proc. Amer. Math. Soc. **134** (2006), no. 2, 465–474.
- [MS] R.A. Macías and C. Segovia, *Lipschitz functions on spaces of homogeneous type*, Adv. in Math. **33** (1979), 257–270.
- [M1] Y. Meyer, *Wavelets and Operators*, translated from the 1990 French original by D.H. Salinger, Cambridge Studies in Advanced Mathematics **37**, Cambridge University Press, Cambridge, 1992.
- [M2] Y. Meyer, *Les nouveaux opérateurs de Calderón–Zygmund*, Astérisque **131** (1985), 237–254.
- [NS1] A. Nagel and E.M. Stein, *On the product theory of singular integrals*, Rev. Mat. Iberoamericana **20** (2004), 531–561.
- [NS2] A. Nagel and E.M. Stein, *The  $\bar{\partial}_b$ -complex on decoupled boundaries in  $\mathbb{C}^n$* , Ann. of Math.(2) **164** (2006), 649–713.

Department of Mathematics, Auburn University, AL 36849-5310, USA.

*E-mail address:* hanyong@auburn.edu

School of Information Technology and Mathematical Sciences, University of South Australia,  
Mawson Lakes SA 5095, Australia,

and

Department of Mathematics, Macquarie University, NSW 2019, Australia.

*E-mail address:* ji.li@mq.edu.au

School of Information Technology and Mathematical Sciences, University of South Australia,  
Mawson Lakes SA 5095, Australia.

*E-mail address:* Lesley.Ward@unisa.edu.au